

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 27, 2002, 14:40:33 ; Search time 35 Seconds  
(without alignments)  
34.264 Million cell updates/sec

Title: US-09-489-760A-1  
Perfect score: 38  
Sequence: 1 VLXDDLEA 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues  
Total number of hits satisfying chosen parameters: 908470

Minimum DB seq Length: 0  
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 1000 summaries

Database :

A\_Geneseq.101002:\*

- 1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*
- 2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*
- 3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*
- 4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*
- 5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*
- 6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*
- 7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*
- 8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*
- 9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*
- 10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*
- 11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*
- 12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*
- 13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*
- 14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*
- 15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*
- 16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*
- 17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*
- 18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*
- 19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*
- 20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*
- 21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*
- 22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*
- 23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	36	94.7	9	20	AAW97374
2	36	94.7	9	20	AAW97375
3	36	94.7	9	20	AAW97376
4	36	94.7	9	20	AAW97377
5	36	94.7	9	20	AAW99195
6	36	94.7	9	20	AAW99196
7	36	94.7	9	20	AAW99197
8	36	94.7	13	20	AAW99198
9	36	94.7	13	20	AAW99199
10	36	94.7	13	20	AAW97414
					AAW97415

11	33	86.8	192	20	AAV34986	Chlamydia pneumoniae
12	33	86.8	226	22	AAU53226	Proionibacterium
13	33	86.8	499	23	ABW90575	Chlamydia pneumoniae
14	32	84.2	4150	21	AAV92707	S. antibioticus 8,
15	31	81.6	149	21	AAW25210	Eucalyptus grandis
16	31	81.6	276	23	ABW48245	Listeria monocytogenes
17	31	81.6	289	23	AAU72987	Neisseria meningitidis
18	31	81.6	289	23	AAU73000	Neisseria meningitidis
19	31	81.6	1042	22	ABW59345	Drosophila melanogaster
20	30	78.9	328	22	AAU86478	Novel human connective
21	30	78.9	984	22	AAU83861	Human colon cancer
22	30	78.9	623	22	AAW96742	Putative P. abyssi
23	30	78.9	1410	21	AAV58574	Sorangenium cellulosum
24	29	76.3	75	22	AAU02010	Gene #37 human secret
25	29	76.3	97	22	AAU74726	Human colon cancer
26	29	76.3	144	21	AAW05154	Arabidopsis thaliana
27	29	76.3	153	21	AAW05154	Arabidopsis thaliana
28	29	76.3	216	23	ABW97149	Novel human diatom
29	29	76.3	222	22	ABW21583	Arabidopsis thaliana
30	29	76.3	224	21	AAW57749	Lactococcus lactis
31	29	76.3	244	23	ABW54117	Arabidopsis thaliana
32	29	76.3	249	21	AAW57748	Arabidopsis thaliana
33	29	76.3	292	22	AAW60860	Streptococcus polydysenteriae
34	29	76.3	297	23	ABW28408	Arabidopsis thaliana
35	29	76.3	358	21	AAW04826	Arabidopsis thaliana
36	29	76.3	358	21	AAW04826	Novel human diatom
37	29	76.3	396	22	ABW29716	Human protein sequ
38	29	76.3	405	22	AAW44128	Arabidopsis thaliana
39	29	76.3	414	21	AAW04825	Arabidopsis thaliana
40	29	76.3	414	21	AAW04825	Arabidopsis thaliana
41	29	76.3	415	21	AAW04824	Arabidopsis thaliana
42	29	76.3	415	21	AAW04824	Arabidopsis thaliana
43	29	76.3	429	22	ABW65976	Drosophila melanogaster
44	29	76.3	468	22	ABW67967	Drosophila melanogaster
45	29	76.3	505	22	ABW67165	Drosophila melanogaster
46	29	76.3	542	21	AAW10764	Human T protein fr
47	29	76.3	547	22	AAW07054	Human gene 4 encod
48	29	76.3	547	22	ABW65102	Human albumin sequ
49	29	76.3	559	22	AAW59765	Human protein sequ
50	29	76.3	565	21	AAW10765	Murine T gene prot
51	29	76.3	717	21	AAW25491	Eucalyptus grandis
52	29	76.3	747	22	ABW63581	Drosophila melanogaster
53	29	76.3	899	22	ABW27691	Novel human diatom
54	29	76.3	932	22	ABW08399	Novel human diatom
55	29	76.3	976	22	AAW56185	Novel human diatom
56	29	76.3	1090	21	AAW85576	Proionibacterium
57	29	76.3	1633	22	AAW23507	HS-ONC-53/1 fragm
58	29	76.3	1837	21	AAW85564	Human EST encoded
59	29	76.3	2473	22	AAW78911	Human homologue of
60	29	76.3	2668	22	AAW78911	Human protein SEQ
61	29	76.3	108	22	AAW66397	Proionibacterium
62	29	76.3	117	22	ABW03818	Human ORF protein
63	29	76.3	195	21	AAW70513	Schistosoma mansoni
64	29	76.3	208	23	ABW90113	Human polypeptide
65	29	76.3	251	17	AAW97012	Influenza virus ma
66	29	76.3	252	21	AAW70054	wild type equine i
67	29	76.3	252	21	AAW70054	Cold-adapted equin
68	29	76.3	252	21	AAW23114	Influenza A virus/
69	29	76.3	252	23	ABW05770	Influenza A/dorm/
70	29	76.3	282	19	AAW34988	Bankia gouldii endo
71	29	76.3	286	22	ABW25477	Novel human diatom
72	29	76.3	293	23	ABW61561	High growth methan
73	29	76.3	293	23	AAW79435	Methylobacterium
74	29	76.3	296	22	AAW38084	Streptococcus pneu
75	29	76.3	296	22	AAW38084	Streptococcus pneu
76	29	76.3	324	22	AAW01080	CPE 83 protein seq
77	29	76.3	324	22	AAW01080	Proionibacterium
78	29	76.3	360	22	AAW93337	Human neuronal apo
79	29	76.3	426	22	ABW02056	Novel human diatom
80	29	76.3	492	22	ABW29153	Novel human diatom
81	29	76.3	550	18	AAW23282	B. lactofermentum
82	29	76.3	550	18	AAW05584	B. lactofermentum
83	29	76.3	550	19	AAW68149	Brevibacterium lac

84	28	73.7	550	19	AAW69551	Brevibacterium lac	157	27	71.1	180	19	AAW98570	H. pylori GHP0 9 p
85	28	73.7	550	19	AAW47398	B. lactofermentum	158	27	71.1	185	22	AAW74588	Human colon cancer
86	28	73.7	550	22	AAW93196	C. glutamicum prote	159	27	71.1	191	23	ABP28217	Streptococcus poly
87	28	73.7	556	22	ABW62959	Drosophila melanog	160	27	71.1	230	21	AAW30259	Arabidopsis thalia
88	28	73.7	638	21	AAW75343	Neisseria gonorrhoe	161	27	71.1	256	21	AAW37157	Arabidopsis thalia
89	28	73.7	638	21	AAW75344	Neisseria meningit	162	27	71.1	256	21	AAW47505	Arabidopsis thalia
90	28	73.7	638	21	AAW75345	Neisseria meningit	163	27	71.1	258	21	AAW52042	Arabidopsis thalia
91	28	73.7	776	23	ABW48568	Listeria monocytog	164	27	71.1	264	21	AAW81568	Streptococcus pneu
92	28	73.7	780	20	AAW41089	Human lectomedin-1	165	27	71.1	274	21	AAW08748	Arabidopsis thalia
93	28	73.7	836	23	AAW15727	Human lectomedin-1	166	27	71.1	276	21	AAW47174	Arabidopsis thalia
94	28	73.7	836	23	AAW15727	Human lectomedin-1	167	27	71.1	279	21	AAW16375	Arabidopsis thalia
95	28	73.7	867	21	AAW75346	Human CLK-2 protei	168	27	71.1	279	21	AAW21798	Arabidopsis thalia
96	28	73.7	867	21	AAW75347	Neisseria gonorrhoe	169	27	71.1	293	22	AAW35037	Arabidopsis thalia
97	28	73.7	867	21	AAW75348	Neisseria meningit	170	27	71.1	293	22	AAW35037	Arabidopsis thalia
98	28	73.7	875	22	AAW02245	Neisseria meningit	171	27	71.1	293	22	AAW35037	Arabidopsis thalia
99	28	73.7	875	22	AAW02245	Neisseria meningit	172	27	71.1	313	19	AAW55983	Arabidopsis thalia
100	28	73.7	901	23	AAW02242	Streptococcus poly	173	27	71.1	317	21	AAW06789	Arabidopsis thalia
101	28	73.7	901	23	AAW02242	Streptococcus poly	174	27	71.1	317	21	AAW06789	Arabidopsis thalia
102	28	73.7	959	14	AAW13523	Sequence of the p1	175	27	71.1	317	21	AAW10941	Arabidopsis thalia
103	28	73.7	959	14	AAW13523	Plasmodium falcipla	176	27	71.1	317	21	AAW10941	Arabidopsis thalia
104	28	73.7	980	21	AAW18294	Domestic mite Bt11	177	27	71.1	317	21	AAW10941	Arabidopsis thalia
105	28	73.7	1017	22	AAW02246	Domestic mite Bt11	178	27	71.1	317	21	AAW10941	Arabidopsis thalia
106	28	73.7	1025	22	AAW02246	Human G protein-co	179	27	71.1	320	21	AAW06788	Arabidopsis thalia
107	28	73.7	1104	22	AAW02246	Novel human diago	180	27	71.1	320	21	AAW06788	Arabidopsis thalia
108	28	73.7	1114	20	AAW41092	Putative P. abyssal	181	27	71.1	320	21	AAW37325	Arabidopsis thalia
109	28	73.7	1114	20	AAW41092	Peptide Seq ID No:	182	27	71.1	320	21	AAW37325	Arabidopsis thalia
110	28	73.7	1123	20	AAW15730	Protein encoded by	183	27	71.1	320	21	AAW37325	Arabidopsis thalia
111	28	73.7	1123	20	AAW15730	Human lectomedin-1	184	27	71.1	324	23	AAW08747	Arabidopsis thalia
112	28	73.7	1123	20	AAW15730	Human lectomedin-1	185	27	71.1	324	23	AAW08747	Arabidopsis thalia
113	28	73.7	1123	20	AAW15730	Human lectomedin-1	186	27	71.1	325	21	AAW22567	Arabidopsis thalia
114	28	73.7	1123	20	AAW15730	Human lectomedin-1	187	27	71.1	325	21	AAW22567	Arabidopsis thalia
115	28	73.7	1320	22	AAW48101	Human lectomedin-1	188	27	71.1	325	21	AAW22567	Arabidopsis thalia
116	28	73.7	1403	20	AAW41087	Propionibacterium	189	27	71.1	327	21	AAW47173	Arabidopsis thalia
117	28	73.7	1403	20	AAW41087	Human lectomedin-1	190	27	71.1	327	21	AAW47173	Arabidopsis thalia
118	28	73.7	1425	22	AAW15725	Human lectomedin-1	191	27	71.1	333	23	ABW04784	Arabidopsis thalia
119	28	73.7	1478	23	AAU10540	Novel human secret	192	27	71.1	338	21	AAW10940	Arabidopsis thalia
120	28	73.7	2175	22	AAW97821	Rat CIRL-2 variant	193	27	71.1	338	21	AAW10940	Arabidopsis thalia
121	28	73.7	3457	15	AAW62504	Human mutated spas	194	27	71.1	350	23	AAW1758	Arabidopsis thalia
122	28	73.7	3457	20	AAW62504	Large polypeptide	195	27	71.1	357	21	AAW10939	Arabidopsis thalia
123	28	73.7	3519	21	AAW92709	Polyprotein encode	196	27	71.1	358	22	ABW18254	Novel human diago
124	28	73.7	3816	21	AAW92708	S. antibioticus 8,	197	27	71.1	381	22	AAW3028	Human protein sequ
125	28	73.7	3829	22	AAW97819	Human spastin prot	198	27	71.1	389	16	AAW08020	Homology vector 52
126	28	73.7	3829	22	AAW97819	Human spastin prot	199	27	71.1	394	23	ABW1822	Prostate cancer-as
127	27	71.1	3830	22	ABW97820	Mouse spastin prot	200	27	71.1	395	21	AAW30258	Arabidopsis thalia
128	27	71.1	63	22	ABW42862	Peptide #10368 enc	201	27	71.1	410	22	ABW69158	Arabidopsis thalia
129	27	71.1	63	22	ABW42862	Human brain expres	202	27	71.1	413	21	AAW06787	Drosophila melanog
130	27	71.1	63	22	ABW42862	Peptide #10712 enc	203	27	71.1	413	21	AAW06787	Arabidopsis thalia
131	27	71.1	63	22	ABW42862	Human ORFX protein	204	27	71.1	414	21	AAW37324	Arabidopsis thalia
132	27	71.1	63	22	ABW42862	Partial mouse CLK-	205	27	71.1	419	22	AAW03534	Human protein kin
133	27	71.1	63	22	ABW42862	Arabidopsis thalia	206	27	71.1	419	22	AAW03534	Human protein kin
134	27	71.1	66	23	ABW11169	Bradyrhizobium jap	207	27	71.1	423	21	AAW14295	Human kinase (PKIN
135	27	71.1	66	23	ABW11169	Arabidopsis thalia	208	27	71.1	423	21	AAW14295	Arabidopsis thalia
136	27	71.1	110	21	AAW05213	Human ORFX protein	209	27	71.1	423	21	AAW14295	Arabidopsis thalia
137	27	71.1	111	22	AAW05213	Human ORFX protein	210	27	71.1	444	23	ABW04731	Novel human diago
138	27	71.1	119	22	AAW05213	Human ORFX protein	211	27	71.1	447	22	AAW03135	Arabidopsis thalia
139	27	71.1	124	21	AAW05213	Human ORFX protein	212	27	71.1	447	22	AAW03135	Arabidopsis thalia
140	27	71.1	127	23	ABW08697	Human PRO polypept	213	27	71.1	463	22	ABW30002	Streptococcus pyog
141	27	71.1	127	23	ABW08697	Human PRO polypept	214	27	71.1	470	21	AAW09117	Streptococcus pyog
142	27	71.1	140	21	AAW99454	Human PRO polypept	215	27	71.1	470	21	AAW09117	Streptococcus pyog
143	27	71.1	140	21	AAW99454	Human PRO polypept	216	27	71.1	474	23	AAW42430	Streptococcus poly
144	27	71.1	140	21	AAW99454	Human PRO polypept	217	27	71.1	477	22	AAW61057	Novel human diago
145	27	71.1	140	21	AAW99454	Human PRO polypept	218	27	71.1	477	22	AAW61057	Novel human diago
146	27	71.1	142	21	AAW42203	Protein of the inv	219	27	71.1	481	23	ABW48428	Arabidopsis thalia
147	27	71.1	158	16	AAW82701	Rice fungal resist	220	27	71.1	519	23	ABW48428	Arabidopsis thalia
148	27	71.1	158	16	AAW82701	Protein encoded by	221	27	71.1	520	12	AAW11820	Streptococcus poly
149	27	71.1	160	23	AAW59944	Human DITP polype	222	27	71.1	531	21	AAW84432	Arabidopsis thalia
150	27	71.1	162	23	AAW59944	Human DITP polype	223	27	71.1	531	21	AAW84432	Arabidopsis thalia
151	27	71.1	172	21	AAW75394	Neisseria gonorrhoe	224	27	71.1	545	22	AAW42429	Arabidopsis thalia
152	27	71.1	172	21	AAW75394	Neisseria gonorrhoe	225	27	71.1	545	22	AAW42429	Arabidopsis thalia
153	27	71.1	172	21	AAW75394	Neisseria meningit	226	27	71.1	574	22	ABW63976	Human polypeptide
154	27	71.1	173	23	ABW30456	Streptococcus poly	227	27	71.1	596	21	AAW42428	Arabidopsis thalia
155	27	71.1	177	22	ABW18253	Novel human diago	228	27	71.1	601	22	ABW60907	Arabidopsis thalia
156	27	71.1	179	21	AAW05211	Arabidopsis thalia	229	27	71.1	621	22	ABW21855	Novel human diago

230	27	71.1	624	23	ABG61779	Novel neurolysin p			
231	27	71.1	637	23	ABR09758	Amino aide sequenc			
232	27	71.1	638	23	AAW47669	MOL9a protein sequ			
233	27	71.1	664	22	AAE04894	Human transporter			
234	27	71.1	681	23	ABR77755	Amino acid sequenc			
235	27	71.1	687	21	AAI69165	A human N-acetylgl			
236	27	71.1	687	21	AAW47670	MOL9b fragment. H			
237	27	71.1	690	23	AAE15988	Human CNGA3 protei			
238	27	71.1	694	23	AAE15987	Human CNGA3 protei			
239	27	71.1	704	22	ABR73485	Human aminopeptida			
240	27	71.1	704	23	ABG61773	Novel neurolysin p			
241	27	71.1	704	23	ABG61774	Novel neurolysin p			
242	27	71.1	704	23	ABG61775	Novel neurolysin p			
243	27	71.1	704	23	ABG61776	Novel neurolysin p			
244	27	71.1	704	23	ABG61777	Novel neurolysin p			
245	27	71.1	704	23	ABG61778	Novel neurolysin p			
246	27	71.1	704	23	ABG61780	Novel neurolysin p			
247	27	71.1	704	23	ABG61781	Novel neurolysin p			
248	27	71.1	704	23	AAG68280	Human NHP protein			
249	27	71.1	704	23	AAW47671	MOL9c protein sequ			
250	27	71.1	709	13	AAW47674	Rabbit endopeptida			
251	27	71.1	714	23	AAU74744	Human protease PRT			
252	27	71.1	757	22	ABG13296	Novel human diagno			
253	27	71.1	765	21	AAV69167	A mature human N-a			
254	27	71.1	765	21	AAV69168	Mouse eps8. Mus m			
255	27	71.1	821	14	AAW34511	Tomato mosaic viru			
256	27	71.1	866	23	ABR04602	Arabidopsis thalia			
257	27	71.1	925	22	AAE06696	Herbicideally activ			
258	27	71.1	1021	23	ABR92803	Novel human diagno			
259	27	71.1	1039	22	ABG15145	M. tuberculosis is			
260	27	71.1	1043	19	AAW57448	Tomato protein SEQ			
261	27	71.1	1083	22	AAW79921	Tobacco mosaic viru			
262	27	71.1	1116	15	AAW60608	Tobacco mosaic vir			
263	27	71.1	1116	23	AAE18345	TMV 126kDa mutant			
264	27	71.1	1116	23	AAE18347	TMV 126kDa mutant			
265	27	71.1	1116	23	AAE19946	TMV 126kDa mutant			
266	27	71.1	1116	23	AAE20136	TMV 126kDa mutant			
267	27	71.1	1116	23	AAE20137	TMV 126kDa mutant			
268	27	71.1	1151	22	ABR04603	Tomato mosaic viru			
269	27	71.1	1151	22	AAW78937	Human protein SEQ			
270	27	71.1	1172	22	AAE18345	Myxobacteria tumo			
271	27	71.1	1346	22	ABR66153	Drosophila melanog			
272	27	71.1	1485	21	AAV69166	A mature human N-a			
273	27	71.1	1616	23	AAE18346	Tobacco mosaic vir			
274	27	71.1	1616	23	AAE18348	TMV 183kDa mutant			
275	27	71.1	1616	23	AAE19947	TMV 183kDa mutant			
276	27	71.1	1616	23	AAE20138	TMV 183kDa mutant			
277	27	71.1	1616	23	ABR04601	Tomato mosaic viru			
278	27	71.1	1616	23	AAE20139	SFY4 non-structura			
279	27	71.1	2431	13	AAW25138	B. forsythus prote			
280	27	71.1	10	20	AAW73901	Escherichia coli s			
281	27	68.4	37	22	AAU09379	Protein #2980 enco			
282	26	68.4	45	22	ABR20981	Human bone marrow			
283	26	68.4	45	22	AAW68745	Peptide #2968 enco			
284	26	68.4	45	22	AAW04286	Human peptidase enco			
285	26	68.4	45	23	ABG38330	Propionibacterium			
286	26	68.4	60	22	AAU48202	Propionibacterium			
287	26	68.4	71	22	AAU57325	Human immune/haema			
288	26	68.4	71	22	AAW90695	Listeria monocytog			
289	26	68.4	77	23	ABR49538	Human OREX protein			
290	26	68.4	82	23	ABP09291	Human OREX protein			
291	26	68.4	82	23	ABP10814	Streptococcus poly			
292	26	68.4	84	23	ABP26553	Streptococcus poly			
293	26	68.4	89	23	ABP26552	Human OREX protein			
294	26	68.4	89	23	ABP03686	Human OREX protein			
295	26	68.4	97	23	ABP03772	Propionibacterium			
296	26	68.4	98	22	AAU54074	Arabidopsis thalia			
297	26	68.4	105	21	AAW12079	Novel human diagno			
298	26	68.4	114	22	ABG10903	RG512 protein RGS			
299	26	68.4	117	20	AAV41025	Rat RGS12 protein			
300	26	68.4	119	20	AAV41003	Human ORF3268 prot			
301	26	68.4	122	21	ABP34295	Arabidopsis thalia			
302	26	68.4	128	21	AAW12078	Arabidopsis thalia			
303	26	68.4	303	26	68.4	140	22	AAW00807	Human bone marrow
304	26	68.4	304	26	68.4	151	21	AAE588251	Lung cancer associ
305	26	68.4	305	26	68.4	157	21	ABR41024	Human OREX ORF788
306	26	68.4	306	26	68.4	157	23	ABP35486	Human synthase-11k
307	26	68.4	307	26	68.4	159	21	ABR41477	Human OREX ORF1241
308	26	68.4	308	26	68.4	159	23	ABP34510	Human OREX ORF1241
309	26	68.4	309	26	68.4	169	22	ABR66997	Human carboxylpept
310	26	68.4	310	26	68.4	173	23	ABP10827	Human OREX protein
311	26	68.4	311	26	68.4	182	22	ABR66835	Purative P. abyssi
312	26	68.4	312	26	68.4	185	21	ABR52827	Eucalyptus grandis
313	26	68.4	313	26	68.4	190	21	AAW54341	Human pancreatic c
314	26	68.4	314	26	68.4	191	20	AAV32126	Human Lysr Interac
315	26	68.4	315	26	68.4	193	21	AAW12077	Arabidopsis thalia
316	26	68.4	316	26	68.4	195	22	ABR69996	Human carboxylpept
317	26	68.4	317	26	68.4	198	22	ABR29159	Novel human diagno
318	26	68.4	318	26	68.4	199	21	AAW60146	Human bone marrow
319	26	68.4	319	26	68.4	202	21	AAW60146	Arabidopsis thalia
320	26	68.4	320	26	68.4	202	22	AAW90239	Corynebacterium gl
321	26	68.4	321	26	68.4	202	22	AAW79267	Propionibacterium
322	26	68.4	322	26	68.4	204	21	AAW42614	Arabidopsis thalia
323	26	68.4	323	26	68.4	204	22	AAW25823	Arabidopsis thalia
324	26	68.4	324	26	68.4	208	21	AAW60145	Arabidopsis thalia
325	26	68.4	325	26	68.4	211	21	AAW60145	Arabidopsis thalia
326	26	68.4	326	26	68.4	212	21	AAW74294	Arabidopsis thalia
327	26	68.4	327	26	68.4	212	21	AAW473322	Arabidopsis thalia
328	26	68.4	328	26	68.4	215	22	AAU077050	Propionibacterium
329	26	68.4	329	26	68.4	217	21	AAW74292	Arabidopsis thalia
330	26	68.4	330	26	68.4	217	22	ABR67335	Arabidopsis thalia
331	26	68.4	331	26	68.4	217	21	AAV74293	Neisseria gonorrhoe
332	26	68.4	332	26	68.4	218	21	AAV74294	Neisseria meningit
333	26	68.4	333	26	68.4	218	21	AAV74295	Neisseria meningit
334	26	68.4	334	26	68.4	220	21	ABR97190	Neisseria gonorrhoe
335	26	68.4	335	26	68.4	220	21	AAW74295	Neisseria meningit
336	26	68.4	336	26	68.4	230	22	AAU57540	Novel human protei
337	26	68.4	337	26	68.4	230	22	ABG29916	Propionibacterium
338	26	68.4	338	26	68.4	233	22	ABG29916	Novel human diagno
339	26	68.4	339	26	68.4	233	22	ABG29916	Arabidopsis thalia
340	26	68.4	340	26	68.4	233	22	ABG29916	Arabidopsis thalia
341	26	68.4	341	26	68.4	233	22	ABG29916	Arabidopsis thalia
342	26	68.4	342	26	68.4	233	22	ABG29916	Arabidopsis thalia
343	26	68.4	343	26	68.4	233	22	ABG29916	Arabidopsis thalia
344	26	68.4	344	26	68.4	233	22	ABG29916	Arabidopsis thalia
345	26	68.4	345	26	68.4	233	22	ABG29916	Arabidopsis thalia
346	26	68.4	346	26	68.4	233	22	ABG29916	Arabidopsis thalia
347	26	68.4	347	26	68.4	233	22	ABG29916	Arabidopsis thalia
348	26	68.4	348	26	68.4	233	22	ABG29916	Arabidopsis thalia
349	26	68.4	349	26	68.4	233	22	ABG29916	Arabidopsis thalia
350	26	68.4	350	26	68.4	233	22	ABG29916	Arabidopsis thalia
351	26	68.4	351	26	68.4	233	22	ABG29916	Arabidopsis thalia
352	26	68.4	352	26	68.4	233	22	ABG29916	Arabidopsis thalia
353	26	68.4	353	26	68.4	233	22	ABG29916	Arabidopsis thalia
354	26	68.4	354	26	68.4	233	22	ABG29916	Arabidopsis thalia
355	26	68.4	355	26	68.4	233	22	ABG29916	Arabidopsis thalia
356	26	68.4	356	26	68.4	233	22	ABG29916	Arabidopsis thalia
357	26	68.4	357	26	68.4	233	22	ABG29916	Arabidopsis thalia
358	26	68.4	358	26	68.4	233	22	ABG29916	Arabidopsis thalia
359	26	68.4	359	26	68.4	233	22	ABG29916	Arabidopsis thalia
360	26	68.4	360	26	68.4	233	22	ABG29916	Arabidopsis thalia
361	26	68.4	361	26	68.4	233	22	ABG29916	Arabidopsis thalia
362	26	68.4	362	26	68.4	233	22	ABG29916	Arabidopsis thalia
363	26	68.4	363	26	68.4	233	22	ABG29916	Arabidopsis thalia
364	26	68.4	364	26	68.4	233	22	ABG29916	Arabidopsis thalia
365	26	68.4	365	26	68.4	233	22	ABG29916	Arabidopsis thalia
366	26	68.4	366	26	68.4	233	22	ABG29916	Arabidopsis thalia
367	26	68.4	367	26					

376	26	68.4	325	23	ABP28691	Streptococcus poly
377	26	68.4	327	21	AAV58618	Protein regulating
378	26	68.4	327	22	AAV58618	C glutamyl prote
379	26	68.4	330	23	ABG66915	Streptococcus pyog
380	26	68.4	330	23	ABG66916	Streptococcus pyog
381	26	68.4	330	23	ABG66917	Streptococcus pyog
382	26	68.4	331	22	AAV66917	C glutamyl prote
383	26	68.4	331	21	AAV66917	Mammalian IGF-1 re
384	26	68.4	333	23	ABG34089	Human Tbx Interact
385	26	68.4	333	23	ABG34127	Human GLUT1 C-term
386	26	68.4	333	23	ABG04778	LDL receptor bindi
387	26	68.4	333	23	ABG04782	LDL receptor bindi
388	26	68.4	333	23	ABG04783	LDL receptor bindi
389	26	68.4	337	23	ABG04785	LDL receptor bindi
390	26	68.4	337	21	AAV52436	Arabidopsis thalia
391	26	68.4	341	21	AAV52435	Arabidopsis thalia
392	26	68.4	342	21	AAV52435	Arabidopsis thalia
393	26	68.4	350	23	ABG66912	Streptococcus pyog
394	26	68.4	350	23	ABG66912	Streptococcus pyog
395	26	68.4	355	22	ABG25843	Corynebacterium gl
396	26	68.4	355	22	ABG25843	Herbically activ
397	26	68.4	357	23	ABG25843	Novel human diago
398	26	68.4	357	23	ABG25843	Novel human diago
399	26	68.4	372	21	AAV39973	Arabidopsis thalia
400	26	68.4	372	21	AAV39973	Arabidopsis thalia
401	26	68.4	405	21	AAV04814	Microspora eve
402	26	68.4	408	21	AAV43289	Human ORFX ORF3053
403	26	68.4	408	21	AAV43289	Arabidopsis thalia
404	26	68.4	408	21	AAV43289	Arabidopsis thalia
405	26	68.4	412	23	ABG77019	Herbically activ
406	26	68.4	412	23	ABG77019	Human protein sequ
407	26	68.4	422	20	AAV79946	Human albumin fusi
408	26	68.4	422	20	AAV79946	Corynebacterium gl
409	26	68.4	424	21	AAV91067	B. forsythus prote
410	26	68.4	427	22	AAV91067	Streptococcus noga
411	26	68.4	434	12	AAV12533	Human polypeptide
412	26	68.4	434	12	AAV12533	Inosine guanosine
413	26	68.4	438	21	AAV84831	Listeria monocytog
414	26	68.4	438	21	AAV84831	A Streptococcus pn
415	26	68.4	452	20	AAV05766	Human PRO216 (vite
416	26	68.4	452	20	AAV05766	Amino acid sequenc
417	26	68.4	452	21	AAV88378	PRO302, vitellogen
418	26	68.4	452	22	AAE06595	Human protein havi
419	26	68.4	452	22	AAE06595	Human PRO302. Hom
420	26	68.4	452	22	AAE06595	Human angiotensin
421	26	68.4	452	22	AAE06595	Human secreted pro
422	26	68.4	452	23	ABG5448	Lactococcus lactis
423	26	68.4	452	23	ABG5448	Streptococcus pneu
424	26	68.4	452	23	ABG5448	Arabidopsis thalia
425	26	68.4	452	23	ABG5448	Arabidopsis thalia
426	26	68.4	452	23	ABG5448	Novel human diago
427	26	68.4	452	23	ABG5448	Arabidopsis thalia
428	26	68.4	452	23	ABG5448	Arabidopsis thalia
429	26	68.4	452	23	ABG5448	Arabidopsis thalia
430	26	68.4	452	23	ABG5448	Arabidopsis thalia
431	26	68.4	452	23	ABG5448	Arabidopsis thalia
432	26	68.4	452	23	ABG5448	Arabidopsis thalia
433	26	68.4	452	23	ABG5448	Arabidopsis thalia
434	26	68.4	452	23	ABG5448	Arabidopsis thalia
435	26	68.4	452	23	ABG5448	Arabidopsis thalia
436	26	68.4	452	23	ABG5448	Arabidopsis thalia
437	26	68.4	452	23	ABG5448	Arabidopsis thalia
438	26	68.4	452	23	ABG5448	Arabidopsis thalia
439	26	68.4	452	23	ABG5448	Arabidopsis thalia
440	26	68.4	452	23	ABG5448	Arabidopsis thalia
441	26	68.4	452	23	ABG5448	Arabidopsis thalia
442	26	68.4	452	23	ABG5448	Arabidopsis thalia
443	26	68.4	452	23	ABG5448	Arabidopsis thalia
444	26	68.4	452	23	ABG5448	Arabidopsis thalia
445	26	68.4	452	23	ABG5448	Arabidopsis thalia
446	26	68.4	452	23	ABG5448	Arabidopsis thalia
447	26	68.4	452	23	ABG5448	Arabidopsis thalia
448	26	68.4	452	23	ABG5448	Arabidopsis thalia



522	26	68.4	1190	22	AAG64164	Human MD36 protein	
523	26	68.4	1190	23	AAU98848	Human MD36 protein	
524	26	68.4	1192	22	AAG64169	Mouse MD36 protein	
525	26	68.4	1212	21	AAG46262	Arabidopsis thalia	
526	26	68.4	1215	21	AAG30842	Arabidopsis thalia	
527	26	68.4	1242	22	ABB63759	Drosophila melanog	
528	26	68.4	1251	22	ABB61254	Drosophila melanog	
529	26	68.4	1352	21	AAG46261	Arabidopsis thalia	
530	26	68.4	1355	21	AAG30841	Arabidopsis thalia	
531	26	68.4	1396	22	ABB29766	Novel human diagno	
532	26	68.4	1405	22	ABB60515	Drosophila melanog	
533	26	68.4	1465	22	AAB31939	Amino acid sequenc	
534	26	68.4	1465	23	ABG61621	Polymeraase III typ	
535	26	68.4	1465	23	ABP26569	Streptococcus poly	
536	26	68.4	1472	22	ABB62283	Drosophila melanog	
537	26	68.4	1502	22	AAM39273	Human polypeptide	
538	26	68.4	1503	19	AAW48845	Human receptor tyr	
539	26	68.4	1511	21	AAG46260	Arabidopsis thalia	
540	26	68.4	1514	23	ABB90837	Herbicidally activ	
541	26	68.4	1565	22	AAW41059	Human polypeptide	
542	26	68.4	1931	22	ABB66948	Drosophila melanog	
543	26	68.4	1982	22	ABG16404	Novel human diagno	
544	26	68.4	1982	22	ABG19656	Novel human diagno	
545	26	68.4	2153	22	AAB83974	Amino acid sequenc	
546	26	68.4	2165	17	AAR94930	RSV RNA-dependent	
547	26	68.4	2165	19	AAW77015	Respiratory syncyt	
548	26	68.4	2165	20	AAW01819	L proteain of Resp	
549	26	68.4	2165	22	AAB31881	Amino acid sequenc	
550	26	68.4	2165	23	ABG67228	Respiratory syncyt	
551	26	68.4	2165	23	ABG67231	Respiratory syncyt	
552	26	68.4	2165	23	ABG67232	Respiratory syncyt	
553	26	68.4	2165	23	ABG67234	Respiratory syncyt	
554	26	68.4	2165	23	ABG67235	Respiratory syncyt	
555	26	68.4	2165	23	ABG67236	Respiratory syncyt	
556	26	68.4	2165	23	ABG67237	Respiratory syncyt	
557	26	68.4	2165	23	ABG67238	Respiratory syncyt	
558	26	68.4	2165	23	ABG67239	Respiratory syncyt	
559	26	68.4	2165	23	ABG67240	Respiratory syncyt	
560	26	68.4	2165	23	ABG67241	Respiratory syncyt	
561	26	68.4	2165	23	ABG67242	Respiratory syncyt	
562	26	68.4	2165	23	ABG67243	Respiratory syncyt	
563	26	68.4	2165	23	ABG67244	Respiratory syncyt	
564	26	68.4	2165	23	ABG67245	Respiratory syncyt	
565	26	68.4	2165	23	ABG67246	Respiratory syncyt	
566	26	68.4	2165	23	ABG67247	Respiratory syncyt	
567	26	68.4	2165	23	ABG67248	Respiratory syncyt	
568	26	68.4	2165	23	ABG67249	Respiratory syncyt	
569	26	68.4	2165	23	ABG67250	Respiratory syncyt	
570	26	68.4	2165	23	ABG67251	Respiratory syncyt	
571	26	68.4	2165	23	ABG67252	Respiratory syncyt	
572	26	68.4	2165	23	ABG67253	Respiratory syncyt	
573	26	68.4	2165	23	ABG67254	Respiratory syncyt	
574	26	68.4	2165	23	ABG67255	Respiratory syncyt	
575	26	68.4	2165	23	ABG67256	Respiratory syncyt	
576	26	68.4	2165	23	ABG67257	Respiratory syncyt	
577	26	68.4	2165	23	ABG67258	Respiratory syncyt	
578	26	68.4	2165	23	ABG67259	Respiratory syncyt	
579	26	68.4	2165	23	ABG67260	Respiratory syncyt	
580	26	68.4	2166	19	AAW48715	RSV isolate 18537	
581	26	68.4	2166	19	AAW48716	RSV vaccine 2B33F	
582	26	68.4	2166	19	AAW48717	RSV vaccine 2B20L	
583	26	68.4	2166	19	AAW48718	RSV revertant 2B33	
584	26	68.4	2166	19	AAW48719	RSV revertant 2B20	
585	26	68.4	2166	19	AAW48714	RSV isolate 2B w11	
586	26	68.4	2166	20	AAW02005	L protein of RSV w	
587	26	68.4	2166	20	AAW02006	L protein of RSV w	
588	26	68.4	2166	20	AAW02007	L protein of RSV v	
589	26	68.4	2166	20	AAW02008	L protein of RSV v	
590	26	68.4	2166	20	AAW02009	L protein of RSV r	
591	26	68.4	2166	20	AAW02010	L protein of RSV r	
592	26	68.4	2594	16	AAW14748	IgG-Fc binding pro	
593	26	68.4	2957	22	ABG22214	Novel human diagno	
594	26	68.4	5405	16	AAW14749	IgG-Fc binding pro	
595	25	65.8		10	22	AAG94234	Human complementar
596	25	65.8		12	22	AAW64577	Human Factor V pro
597	25	65.8		15	23	ABW07889	Phosphoenolpyruvat
598	25	65.8		22	20	AAV19066	Lectihin:cholester
599	25	65.8		22	20	AAV18812	Lectihin:cholester
600	25	65.8		22	20	AAV18549	Lectihin:cholester
601	25	65.8		22	20	AAV19340	Lectihin:cholester
602	25	65.8		43	22	ABB29797	Peptide #2448 enco
603	25	65.8		43	22	ABB34975	Peptide #2481 enco
604	25	65.8		43	22	ABB20383	Human brain expres
605	25	65.8		43	22	AAW57782	Human brain expres
606	25	65.8		43	22	AAW68157	Human bone marrow
607	25	65.8		43	22	AAW15980	Peptide #2414 enco
608	25	65.8		43	22	AAW28480	Peptide #2517 enco
609	25	65.8		43	22	AAW03715	Peptide #2397 enco
610	25	65.8		46	22	ABB39742	Peptide #7248 enco
611	25	65.8		46	22	ABB24382	Protein #6381 enco
612	25	65.8		46	22	AAW60462	Human brain expres
613	25	65.8		46	22	AAW73106	Human bone marrow
614	25	65.8		46	22	AAW19842	Human bone marrow
615	25	65.8		46	22	AAW33324	Peptide #6276 enco
616	25	65.8		46	22	ABG42952	Peptide #7361 enco
617	25	65.8		46	23	ABG42952	Human pepide enco
618	25	65.8		46	23	AAU21027	Human novel foetal
619	25	65.8		54	22	AAW93180	Human digestive sy
620	25	65.8		54	22	AAU20115	Human liver associ
621	25	65.8		54	23	ABP40976	Human liver antige
622	25	65.8		62	21	AAW56702	Human prostate can
623	25	65.8		62	21	ABW44024	Peptide #11530 enc
624	25	65.8		62	22	ABB26928	Protein #8927 enco
625	25	65.8		62	22	AAW65040	Human brain expres
626	25	65.8		62	22	AAW7754	Human bone marrow
627	25	65.8		62	22	AAW7754	Human bone marrow
628	25	65.8		63	22	AAW37971	Peptide #12008 enc
629	25	65.8		75	22	AAW61935	Human OREF protein
630	25	65.8		82	21	AAW64392	Protonibacterium
631	25	65.8		82	21	AAW36683	Human secreted pro
632	25	65.8		83	21	AAW89890	Aldehyde/ketone re
633	25	65.8		89	21	AAW40420	Human OREF ORF184
634	25	65.8		89	23	ABP03964	Human OREF protein
635	25	65.8		92	21	AAW91959	Human cytoskeleton
636	25	65.8		97	21	AAW02750	Human secreted pro
637	25	65.8		98	21	AAW08473	Arabidopsis thalia
638	25	65.8		98	21	AAW10752	Arabidopsis thalia
639	25	65.8		98	21	AAW46765	Arabidopsis thalia
640	25	65.8		98	21	AAW46778	Arabidopsis thalia
641	25	65.8		98	21	AAW61051	Arabidopsis thalia
642	25	65.8		103	21	AAW52524	Arabidopsis thalia
643	25	65.8		105	22	AAU45612	Protonibacterium
644	25	65.8		107	23	ABP29188	Streptococcus poly
645	25	65.8		109	23	ABP01132	Human OREF protein
646	25	65.8		117	19	AAW61201	Streptococcus pneu
647	25	65.8		117	23	ABP54620	S. pneumoniae Sp07
648	25	65.8		121	22	AAW98399	Escherichia coli P
649	25	65.8		125	22	AAW95744	Human reproductive
650	25	65.8		127	22	AAU15097	Protein encoded by
651	25	65.8		131	22	ABG22426	Novel human diagno
652	25	65.8		132	21	AAW25233	Arabidopsis thalia
653	25	65.8		133	22	AAW40277	Protonibacterium
654	25	65.8		138	22	ABG16867	Novel human diagno
655	25	65.8		146	22	ABG15095	Novel human diagno
656	25	65.8		150	21	AAW05388	Arabidopsis thalia
657	25	65.8		153	10	AAW93718	Adenine-phosphorib
658	25	65.8		154	19	AAW38636	S. pneumoniae putr
659	25	65.8		156	22	ABG12208	Novel human diagno
660	25	65.8		157	22	AAW53065	Protonibacterium
661	25	65.8		164	14	AAW43008	Novel human diagno
662	25	65.8		168	21	AAW41433	Arabidopsis thalia
663	25	65.8		170	21	AAW90791	M. jannaschii M016
664	25	65.8		170	21	AAW90792	P. fultuosus PF00F3
665	25	65.8		170	21	AAW90793	M. thermophilus
666	25	65.8		170	21	AAW90794	P. horikoshii PHBNO
667	25	65.8		173	22	AAW4184	Protonibacterium
668	25	65.8		175	22	AAW67178	Protonibacterium

668	25	65.8	175	22	ABG07908	Novel human diagno	741	25	65.8	283	23	ABB54449	Lactococcus lactis
669	25	65.8	176	21	AA643709	Arabidopsis thalia	742	25	65.8	285	20	AAV35437	Chlamydia pneumonia
670	25	65.8	181	19	AAM55093	Streptococcus pneu	743	25	65.8	286	21	AA631782	Arabidopsis thalia
671	25	65.8	181	23	ABP54587	S. pneumoniae sp04	744	25	65.8	286	21	AA639245	Arabidopsis thalia
672	25	65.8	185	22	AB63757	Human prostate can	745	25	65.8	287	21	AA606373	Arabidopsis thalia
673	25	65.8	195	21	AA631784	Arabidopsis thalia	746	25	65.8	287	21	AA606373	Arabidopsis thalia
674	25	65.8	198	20	AA632174	Chlamydia psittaci	747	25	65.8	287	21	AA640535	Arabidopsis thalia
675	25	65.8	202	21	AA64551	Arabidopsis thalia	748	25	65.8	287	21	AA643160	Arabidopsis thalia
676	25	65.8	202	21	AA640537	Arabidopsis thalia	749	25	65.8	289	21	AA604559	Arabidopsis thalia
677	25	65.8	204	16	AA659481	Prolactin peptide	750	25	65.8	289	23	ABB81658	Arabidopsis thalia
678	25	65.8	205	23	AB694341	Chlamydia trachoma	751	25	65.8	291	22	ABB83270	Human FNI263 prot
679	25	65.8	207	22	AB659084	Drosophila melanog	752	25	65.8	292	21	AA652495	Drosophila melanog
680	25	65.8	207	22	AB617353	Novel human diagno	753	25	65.8	292	21	AA655790	Arabidopsis thalia
681	25	65.8	211	22	AA663794	Human prostate can	754	25	65.8	293	21	AA656955	Arabidopsis thalia
682	25	65.8	212	21	AA657741	Arabidopsis thalia	755	25	65.8	294	22	AA659244	Arabidopsis thalia
683	25	65.8	215	20	AA635572	Protein involved i	756	25	65.8	295	23	ABP39621	Arabidopsis thalia
684	25	65.8	217	21	AA641432	Arabidopsis thalia	757	25	65.8	301	21	AB655841	Drosophila melanog
685	25	65.8	219	21	AA618104	Pinus radiata prot	758	25	65.8	298	22	AAU97736	Stephylococcus epi
686	25	65.8	221	21	AB607850	Amno acid sequenc	759	25	65.8	300	22	AB607888	Drosophila melanog
687	25	65.8	221	23	AB692471	Herbicidally activ	760	25	65.8	301	21	AA619759	Phosphoenolpyruvat
688	25	65.8	222	21	AA618186	Streptococcus pneu	761	25	65.8	301	21	AA619858	Arabidopsis thalia
689	25	65.8	222	21	AA618186	Streptococcus pneu	762	25	65.8	301	21	AA619858	Arabidopsis thalia
690	25	65.8	222	22	AB610582	Novel human diagno	763	25	65.8	301	21	AA619858	Arabidopsis thalia
691	25	65.8	223	22	AB662225	T. gondii max glutat	764	25	65.8	302	21	AA650233	Arabidopsis thalia
692	25	65.8	230	20	AA629015	T. gondii immunoge	765	25	65.8	303	21	AA650935	Arabidopsis thalia
693	25	65.8	230	21	AA656804	Arabidopsis thalia	766	25	65.8	308	21	AA619857	Arabidopsis thalia
694	25	65.8	230	22	AAU25486	T. gondii immunoge	767	25	65.8	308	23	AAU98409	Arabidopsis thalia
695	25	65.8	232	22	AAU55593	Propionibacterium	768	25	65.8	309	21	AA625440	Propionibacterium
696	25	65.8	235	22	AA695578	Human protein sequ	769	25	65.8	310	21	AA619856	Arabidopsis thalia
697	25	65.8	236	23	AB649891	Listeria monocylog	770	25	65.8	310	21	AA619856	Arabidopsis thalia
698	25	65.8	236	23	AB697746	T. maritima RNaseH	771	25	65.8	312	21	AA625489	Arabidopsis thalia
699	25	65.8	238	23	AB609140	Thermotolerant rib	772	25	65.8	313	21	AA650933	Arabidopsis thalia
700	25	65.8	239	20	AA686011	Klebsiella pneumon	773	25	65.8	315	22	AB663576	Arabidopsis thalia
701	25	65.8	242	22	AAU36101	Arabidopsis thalia	774	25	65.8	315	22	AAU34549	Drosophila melanog
702	25	65.8	243	21	AA626271	Human polypeptide	775	25	65.8	316	21	AA637740	E. coli cellular p
703	25	65.8	244	22	AA640345	Arabidopsis thalia	776	25	65.8	319	21	AA606372	Arabidopsis thalia
704	25	65.8	248	21	AA633941	Arabidopsis thalia	777	25	65.8	319	21	AA633159	Arabidopsis thalia
705	25	65.8	249	21	AA604530	Arabidopsis thalia	778	25	65.8	319	22	AAU33513	Arabidopsis thalia
706	25	65.8	249	21	AA640536	Arabidopsis thalia	779	25	65.8	319	22	AA628677	Enterococcus faeca
707	25	65.8	250	23	AAE20496	Streptococcus muta	780	25	65.8	320	22	AAU14247	S. epidermidis ope
708	25	65.8	250	23	AAE20586	Neisseria meningit	781	25	65.8	321	22	AAU38385	Human novel protei
709	25	65.8	252	21	AA674831	Neisseria meningit	782	25	65.8	322	22	AAU36438	Salmonella typhi c
710	25	65.8	252	21	AA674832	Corynebacterium gl	783	25	65.8	323	22	AAU38141	Pseudomonas aerugi
711	25	65.8	252	22	AA690689	Arabidopsis thalia	784	25	65.8	323	22	AAU38141	Salmonella typhi c
712	25	65.8	252	22	AA676512	Arabidopsis thalia	785	25	65.8	325	22	AA621873	Isoprenoid related
713	25	65.8	254	21	AA652497	Novel human diagno	786	25	65.8	325	22	AA690796	C. glutamicum prote
714	25	65.8	254	22	ABG09817	Novel human diagno	787	25	65.8	325	22	AA680081	Corynebacterium gl
715	25	65.8	254	22	ABG12207	Novel human diagno	788	25	65.8	327	22	AAU35303	Arabidopsis thalia
716	25	65.8	255	21	AA652496	Arabidopsis thalia	789	25	65.8	327	22	AAU35303	Enterococcus faeca
717	25	65.8	255	21	AA652496	Arabidopsis thalia	790	25	65.8	327	22	AAU35303	Human protein rela
718	25	65.8	257	22	ABG15097	Novel human diagno	791	25	65.8	328	21	AA607255	Arabidopsis thalia
719	25	65.8	258	13	AA622932	Envelope proteins	792	25	65.8	328	21	AA619768	Arabidopsis thalia
720	25	65.8	258	23	AAE23874	Human protein rela	793	25	65.8	332	21	AA619768	Arabidopsis thalia
721	25	65.8	263	22	AAE23874	Human Rf42 protein	794	25	65.8	332	21	AA650232	Arabidopsis thalia
722	25	65.8	263	22	AAE23874	Novel human diagno	795	25	65.8	332	21	AA650232	Drosophila melanog
723	25	65.8	264	22	AA681596	S. epidermidis ope	796	25	65.8	333	23	AB646157	Arabidopsis thalia
724	25	65.8	265	21	AA631783	Bacteriophage DP-1	797	25	65.8	336	21	AB604781	Arabidopsis thalia
725	25	65.8	268	21	AA616700	Streptococcus pneu	798	25	65.8	336	21	AA619767	LDL receptor bindi
726	25	65.8	275	22	AAU37848	Arabidopsis thalia	799	25	65.8	337	21	AA607254	Arabidopsis thalia
727	25	65.8	276	21	AA643708	Arabidopsis thalia	800	25	65.8	337	23	AB692086	Arabidopsis thalia
728	25	65.8	276	23	AB677642	Arabidopsis thalia	801	25	65.8	337	23	AB692086	Herbicidally activ
729	25	65.8	279	21	AA606374	AMEPV NTP pyrophos	802	25	65.8	339	21	AA658297	Listeria monocylog
730	25	65.8	279	21	AA606374	Arabidopsis thalia	803	25	65.8	340	21	AA658297	Lung cancer associ
731	25	65.8	279	21	AA606374	Arabidopsis thalia	804	25	65.8	342	21	AA626770	Arabidopsis thalia
732	25	65.8	279	21	AA606374	Arabidopsis thalia	805	25	65.8	343	21	AA626770	Arabidopsis thalia
733	25	65.8	280	22	AAU87645	Novel central nerv	806	25	65.8	345	21	AA638499	Amino acid sequenc
734	25	65.8	280	22	AAU87645	Novel central nerv	807	25	65.8	352	22	AA638499	Arabidopsis thalia
735	25	65.8	280	22	AAU87645	Novel central nerv	808	25	65.8	352	22	AA638499	Putative P. abyssi
736	25	65.8	280	22	AAU87645	Human ovariian and/	809	25	65.8	354	21	AA638499	Listeria monocylog
737	25	65.8	280	22	AAU87645	Human ovariian and/	810	25	65.8	354	21	AA638499	Streptococcus pneu
738	25	65.8	280	22	AAU87645	Human ovariian and/	811	25	65.8	358	22	AAU23072	Human protein c pr
739	25	65.8	283	22	AAU87645	Human ovariian anti	812	25	65.8	361	22	AA622401	Novel human enzym
740	25	65.8	283	22	AAU87645	Novel human diagno	813	25	65.8	362	22	AA622401	Novel human diagno
741	25	65.8	283	22	AAU87645	Novel human diagno	814	25	65.8	367	10	AA622401	Ret. d gene product
742	25	65.8	283	22	AAU87645	Novel human diagno	815	25	65.8	367	10	AA622401	Human HMG protein

814	25	65.8	367	23	ABBS0099	Listeria monocytog	887	25	65.8	419	23	AAU99014	Human Protein C zy
815	25	65.8	368	13	AAAR2092	Protein "e" encode	888	25	65.8	419	23	AAU99015	Human Protein C zy
816	25	65.8	368	18	AAAW14835	Protein encoded by	889	25	65.8	419	23	AAU99016	Human Protein C zy
817	25	65.8	368	22	ABBS63805	Drosophila melanog	890	25	65.8	419	23	AAU99017	Human Protein C zy
818	25	65.8	368	22	AAU34140	Staphylococcus aur	891	25	65.8	419	23	AAU99018	Human Protein C zy
819	25	65.8	368	22	AAU36800	Staphylococcus aur	892	25	65.8	419	23	AAU99019	Human Protein C zy
820	25	65.8	369	23	ABP2311	Streptococcus poly	893	25	65.8	419	23	AAU99020	Human Protein C zy
821	25	65.8	373	20	AAU37196	Protein Involved I	894	25	65.8	419	23	AAU99021	Human Protein C zy
822	25	65.8	374	22	AAE03659	Human SFOP TRAF-1	895	25	65.8	419	23	AAU99022	Human Protein C zy
823	25	65.8	374	22	AAAB90799	Human shear stress	896	25	65.8	419	23	AAU99023	Human Protein C zy
824	25	65.8	375	21	AAAG10023	Arabidopsis thalia	897	25	65.8	419	23	AAU99024	Human Protein C zy
825	25	65.8	375	21	AAAG27623	Arabidopsis thalia	898	25	65.8	419	23	AAU99025	Human Protein C zy
826	25	65.8	376	22	AAU38341	Salmonella typhi c	899	25	65.8	419	23	AAU99026	Human Protein C zy
827	25	65.8	377	21	AAAG21649	Arabidopsis thalia	900	25	65.8	419	23	AAU99027	Human Protein C zy
828	25	65.8	377	21	AAAG52572	Arabidopsis thalia	901	25	65.8	419	23	AAU99028	Human Protein C zy
829	25	65.8	377	22	AAAB64759	Drosophila melanog	902	25	65.8	419	23	AAU99029	Human Protein C zy
830	25	65.8	379	21	AAAG07253	Arabidopsis thalia	903	25	65.8	419	23	AAU99030	Human Protein C zy
831	25	65.8	380	22	ABB71363	Drosophila melanog	904	25	65.8	419	23	AAU99031	Human Protein C zy
832	25	65.8	380	23	ABP27310	Streptococcus poly	905	25	65.8	419	23	AAU99032	Human Protein C zy
833	25	65.8	381	23	ABP25733	Streptococcus poly	906	25	65.8	419	23	AAU99033	Human Protein C zy
834	25	65.8	382	22	ABBA4524	Human GPCR2 polype	907	25	65.8	419	23	AAU99034	Human Protein C zy
835	25	65.8	385	21	AAAB58286	Lung cancer associ	908	25	65.8	419	23	AAU99035	Human Protein C zy
836	25	65.8	395	21	AAAG13429	Arabidopsis thalia	909	25	65.8	419	23	AAU99036	Human Protein C zy
837	25	65.8	395	21	AAAG37835	Arabidopsis thalia	910	25	65.8	419	23	AAU99037	Human Protein C zy
838	25	65.8	397	21	AAAG56954	Arabidopsis thalia	911	25	65.8	419	23	AAU99038	Human Protein C zy
839	25	65.8	397	21	AAAG59243	Arabidopsis thalia	912	25	65.8	419	23	AAU99039	Human Protein C zy
840	25	65.8	401	20	AAV17219	Human secreted pro	913	25	65.8	419	23	AAU99040	Human Protein C zy
841	25	65.8	401	22	ABG04175	Novel human diagno	914	25	65.8	419	23	AAU99041	Human Protein C zy
842	25	65.8	401	22	ABG04196	Novel human diagno	915	25	65.8	419	23	AAU99042	Human Protein C zy
843	25	65.8	401	22	AAU38989	Human secreted pro	916	25	65.8	419	23	AAU99043	Human Protein C zy
844	25	65.8	401	22	AAAB93152	Human protein sequ	917	25	65.8	419	23	AAU99044	Human Protein C zy
845	25	65.8	401	23	ABBS5698	Human polypeptide	918	25	65.8	419	23	AAU99045	Human Protein C zy
846	25	65.8	403	23	AAO14999	Rhodotorula minuta	919	25	65.8	419	23	AAU99046	Human Protein C zy
847	25	65.8	404	20	AAV34540	Porphorymonas ging	920	25	65.8	419	23	AAU99047	Human Protein C zy
848	25	65.8	405	22	ABBS59412	Drosophila melanog	921	25	65.8	419	23	AAU99048	Human Protein C zy
849	25	65.8	405	22	AAAB98956	Murine CALP prote	922	25	65.8	419	23	AAU99049	Human Protein C zy
850	25	65.8	408	21	AAAG38498	Arabidopsis thalia	923	25	65.8	419	23	AAU99050	Human Protein C zy
851	25	65.8	411	23	ABBB90946	Herbicidally activ	924	25	65.8	419	23	AAU99051	Human Protein C zy
852	25	65.8	412	21	AAAG56953	Arabidopsis thalia	925	25	65.8	419	23	AAU99052	Human Protein C zy
853	25	65.8	412	22	AAU087165	Novel central nerv	926	25	65.8	419	23	AAU99053	Human Protein C zy
854	25	65.8	413	23	ABP27412	Streptococcus poly	927	25	65.8	419	23	AAU99054	Human Protein C zy
855	25	65.8	415	21	AAV56803	Truncated human pr	928	25	65.8	419	23	AAU99055	Human Protein C zy
856	25	65.8	416	22	AAAM25679	Human protein sequ	929	25	65.8	419	23	AAU99056	Human Protein C zy
857	25	65.8	419	14	AAAR35760	Human protein C (PC) . H	930	25	65.8	419	23	AAU99057	Human Protein C zy
858	25	65.8	419	19	AAAW27253	Primary structure	931	25	65.8	419	23	AAU99058	Human Protein C zy
859	25	65.8	419	21	AAAG22185	Arabidopsis thalia	932	25	65.8	419	23	AAU99059	Human Protein C zy
860	25	65.8	419	21	AAAG48520	Arabidopsis thalia	933	25	65.8	419	23	AAU99060	Human Protein C zy
861	25	65.8	419	22	AAE08625	Human mature wild	934	25	65.8	419	23	AAU99061	Human Protein C zy
862	25	65.8	419	22	AAE08627	Human protein C de	935	25	65.8	419	23	AAU99062	Human Protein C zy
863	25	65.8	419	22	AAE08628	Human protein C de	936	25	65.8	419	23	AAU99063	Human Protein C zy
864	25	65.8	419	22	AAE08629	Human protein C de	937	25	65.8	419	23	AAU99064	Human Protein C zy
865	25	65.8	419	22	AAE08630	Human protein C de	938	25	65.8	419	23	AAU99065	Human Protein C zy
866	25	65.8	419	22	AAE08631	Human protein C de	939	25	65.8	419	23	AAU99066	Human Protein C zy
867	25	65.8	419	22	AAE08632	Human protein C de	940	25	65.8	419	23	AAU99067	Human Protein C zy
868	25	65.8	419	22	AAE08633	Human protein C de	941	25	65.8	419	23	AAU99068	Human Protein C zy
869	25	65.8	419	22	AAE08634	Human protein C de	942	25	65.8	419	23	AAU99069	Human Protein C zy
870	25	65.8	419	22	AAE08635	Human protein C de	943	25	65.8	419	23	AAU99070	Human Protein C zy
871	25	65.8	419	22	AAE08636	Human protein C de	944	25	65.8	419	23	AAU99071	Human Protein C zy
872	25	65.8	419	22	AAE08637	Human protein C de	945	25	65.8	419	23	AAU99072	Human Protein C zy
873	25	65.8	419	22	AAE08638	Human protein C de	946	25	65.8	419	23	AAU99073	Human Protein C zy
874	25	65.8	419	22	AAE08639	Human protein C de	947	25	65.8	419	23	AAU99074	Human Protein C zy
875	25	65.8	419	22	AAE08640	Human protein C de	948	25	65.8	419	23	AAU99075	Human Protein C zy
876	25	65.8	419	22	AAE08641	Human protein C de	949	25	65.8	419	23	AAU99076	Human Protein C zy
877	25	65.8	419	22	AAE08642	Human protein C de	950	25	65.8	419	23	AAU99077	Human Protein C zy
878	25	65.8	419	22	AAE08643	Human protein C de	951	25	65.8	419	23	AAU99078	Human Protein C zy
879	25	65.8	419	22	AAE08644	Human protein C de	952	25	65.8	419	23	AAU99079	Human Protein C zy
880	25	65.8	419	22	AAE08645	Human protein C de	953	25	65.8	419	23	AAU99080	Human Protein C zy
881	25	65.8	419	22	AAE08646	Human protein C de	954	25	65.8	419	23	AAU99081	Human Protein C zy
882	25	65.8	419	22	AAE08647	Human protein C de	955	25	65.8	419	23	AAU99082	Human Protein C zy
883	25	65.8	419	22	AAE08648	Human protein C de	956	25	65.8	419	23	AAU99083	Human Protein C zy
884	25	65.8	419	22	AAE08649	Human protein C de	957	25	65.8	419	23	AAU99084	Human Protein C zy
885	25	65.8	419	22	AAE08650	Human protein C de	958	25	65.8	419	23	AAU99085	Human Protein C zy
886	25	65.8	419	22	AAE08651	Human protein C de	959	25	65.8	419	23	AAU99086	Human Protein C zy

960	25	65.8	419	23	AAU99087	Human Protein C zy
961	25	65.8	419	23	AAU99088	Human Protein C zy
962	25	65.8	419	23	AAU99089	Human Protein C zy
963	25	65.8	419	23	AAU99090	Human Protein C zy
964	25	65.8	419	23	AAU99091	Human Protein C zy
965	25	65.8	419	23	AAU99092	Human Protein C zy
966	25	65.8	419	23	AAU99093	Human Protein C zy
967	25	65.8	419	23	AAU99094	Human Protein C zy
968	25	65.8	419	23	AAU99095	Human Protein C zy
969	25	65.8	419	23	AAU99096	Human Protein C zy
970	25	65.8	419	23	AAU99097	Human Protein C zy
971	25	65.8	425	23	AAU99098	Human Protein C zy
972	25	65.8	427	21	AAH15967	Rhodotorula minuta
973	25	65.8	427	22	AAU29342	E. coli proliferat
974	25	65.8	431	22	ABB60316	Novel mar regulat
975	25	65.8	433	21	AAU17522	Drosophila melanog
976	25	65.8	440	22	AAU74623	Saitodopsis thalia
977	25	65.8	441	20	AAU36680	Saitodopsis thalia
978	25	65.8	444	20	ABP41858	Chlamydia trachoma
979	25	65.8	446	21	AAU38497	Human ovarian anti
980	25	65.8	449	23	ABP40828	Arabidopsis thalia
981	25	65.8	451	23	ABB83793	Staphylococcus epi
982	25	65.8	453	22	AAU40279	Fungal decaprenyl
983	25	65.8	456	22	ABG07909	Human polypeptide
984	25	65.8	458	23	ABR77059	Novel human diagno
985	25	65.8	460	9	AAH1104	Human protein sequ
986	25	65.8	460	12	AAU13537	Sequence of human
987	25	65.8	460	12	AAU13538	Human Protein C zy
988	25	65.8	460	12	AAU13523	Human Protein C zy
989	25	65.8	460	18	AAU25086	Human Protein C zy
990	25	65.8	461	7	AAU60001	Human protein C
991	25	65.8	461	9	AAU70855	Sequence of polype
992	25	65.8	461	9	AAH10205	Human Protein C.
993	25	65.8	461	10	AAU90070	Human protein C.
994	25	65.8	461	10	AAU90070	Human protein C.
995	25	65.8	461	10	AAU90414	Zymogen form of hu
996	25	65.8	461	12	AAU13074	Hybrid protein of
997	25	65.8	461	12	AAU13081	Protein C precursor
998	25	65.8	461	12	AAU13539	Human protein C.
999	25	65.8	461	12	AAU13540	Human Protein C zy
1000	25	65.8	461	12	AAU13622	Human protein C.

## ALIGNMENTS

XX	AA997374	AA997374	standard; Protein: 9 AA.
XX	AA997374	AA997374	
XX	13-MAY-1999	(first entry)	
XX	HA-1	R-allele sequence.	
XX	Intron: minor histocompatibility antigen HA-1: typing allele:		
KW	H allele: R allele: polymorphic nucleotide; HA-1 typing;		
KW	bone marrow transplant; severe aplastic anaemia; leukaemia;		
XX	immune deficiency disease; ss.		
XX	Homo sapiens.		
OS	MO9905313-A2.		
PN	04-FEB-1999.		
PD	23-JUL-1998:	98MO-EF04928.	
XX	02-JUN-1998:	98EP-0870125.	
PR	23-JUL-1997:	97EP-0202303.	
XX	(UYLE-) RICKSONIV LEIDEN.		

XX	Goulmy E;
PI	
XX	WPI; 1999-142960/12.
DR	
XX	
PT	Typing minor histocompatibility antigen HA-1 - by amplifying and
PI	identifying HA-1 R and HA-1 H alleles of HA-1, used for, e.g.
PT	detection of genetic aberrances
XX	
XX	Claim 13; Fig 5; 59pp; English.
XX	
CC	The present sequence represents part of the minor histocompatibility
CC	antigen HA-1 R-allele. The specification describes methods for typing
CC	alleles (preferably the H and R alleles) of the minor histocompatibility
CC	antigen HA-1 in a sample, which comprise detecting polymorphic
CC	nucleotides in the cDNA or genomic nucleic acids of the alleles. The
CC	methods can be used for HA-1 typing for bone marrow transplants, severe
CC	aplastic anaemia, leukaemia and immune deficiency diseases, as well as
CC	detection of genetic aberrances. The probes and primers of the invention
CC	can be used to screen for the HA-1 alleles. The HA-1 peptides can be
CC	used anti-idiotypic B cells and/or T cells and antibodies.
XX	
XX	
SO	Sequence 9 AA:
	Query Match
	Best Local Similarity 94.7%; Score 36; DB 20; length 9;
	Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0
OY	1 VLXDDLLFA 9
DB	1 VLKDDLLFA 9

XX	RESULT 2
XX	AAW97375
ID	AAW97375 standard; Protein; 9 AA.
XX	
AC	AAW97375;
XX	
DT	13-MAY-1999 (first entry)
XX	
DE	HA-1 H-allele sequence.
XX	
XX	Ittron; minor histocompatibility antigen HA-1; typing allele;
KW	H allele; R allele; polymorphic nucleotide; HA-1 typing;
KW	Bone marrow transplant; severe aplastic anaemia; leukaemia;
XX	Immune deficiency disease; ss.
XX	
OS	Homo sapiens.
XX	
PN	W0905313-A2.
XX	
PD	04-FEB-1999.
XX	
PF	23-JUL-1998; 98W0-EP04928.
XX	
PR	02-JUN-1998; 98EP-0870125.
PR	23-JUL-1997; 97EP-0202303.
XX	
PA	(UYLE-) RIJXSUNIV LEIDEN.
XX	
PT	Goulmy E;
XX	
DR	WPI; 1999-142960/12.
XX	
PT	Typing minor histocompatibility antigen HA-1 - by amplifying and
PT	identifying HA-1 R and HA-1 H alleles of HA-1, used for, e.g.
XX	detection of genetic aberrances
PS	Claim 18; Fig 5; 5pp; English.
XX	
CC	The present sequence represents part of the minor histocompatibility
CC	antigen HA-1 H-allele. The specification describes methods for typin

alleles (preferably the H and R alleles) of the minor histocompatibility antigen HA-1 in a sample, which comprise detecting polymorphic nucleotides in the CDNA or genomic nucleic acids of the alleles. The methods can be used for HA-1 typing for bone marrow transplants, severe aplastic anaemia, leukaemia and immune deficiency diseases, as well as detection of genetic aberrances. The probes and primers of the invention can be used to screen for the HA-1 alleles. The HA-1 peptides can be used anti-idiotypic B cells and/or T cells and antibodies.

Sequence 9 AA:

Query Match 94.7%; Score 36; DB 20; Length 9;  
Best Local Similarity 88.9%; Pred. No. 7.8e+05;

Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 VLXDDLLEA 9  
1 VLHDDLLEA 9

RESULT 3  
AAW97572  
ID AAW97572 standard; peptide; 9 AA.

XX AAW97572;

AC 20-MAY-1999 (first entry)

XX T-cell epitope from the minor histocompatibility antigen HA-1.

XX T-cell epitope; minor histocompatibility antigen HA-1; vaccine;

KW transplant rejection; Graft-versus-Host Disease; autoimmune disease;

KW neoplastic haematopoietic cell.

XX Homo sapiens.

XX Key Location/Qualifiers

XX MISC-difference 3 /note="His or Arg"

XX WO9905173-A1.

XX 04-FEB-1999.

XX 23-JUL-1998; 98WO-NL00424.

XX 23-JUL-1997; 97EP-0202303.

XX (UYLE-) RIJXSUNIV LEIDEN.

XX Engelhard VH, Goulmy EAJM, Hunt DF;

XX WPI, 1999-142855/12.

XX Immunogenic peptide from minor histocompatibility antigen HA-1 -

XX useful for inducing tolerance to transplants and prevent rejection

XX or graft-versus-host disease

XX Claim 1; Page 39; 57pp; English.

XX The present sequence represents an immunogenic peptide constituting a T-cell epitope, obtainable from the minor histocompatibility antigen HA-1. The peptide can be used in vaccines or pharmaceutical formulations as medicines to induce tolerance for transplants, so as to prevent rejection and/or graft-versus-Host Disease, or to treat autoimmune diseases. Neoplastic haematopoietic cells presenting the peptides, in an HLA class I context, can be eliminated after specific recognition of the peptides. The peptides can also be used to raise antibodies, T-cell receptor, B- and T-cells.

XX Sequence 9 AA:

Query Match 94.7%; Score 36; DB 20; Length 9;

Best Local Similarity 100.0%; Pred. No. 7.8e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VLXDDLLEA 9  
1 VLHDDLLEA 9

RESULT 4  
AAW99195  
ID AAW99195 standard; peptide; 9 AA.

XX AAW99195;

AC 20-MAY-1999 (first entry)

XX Minor histocompatibility antigen HA-1 T-cell epitope #1.

KW Minor histocompatibility antigen; HA-1; T-cell epitope; immunological;

KW graft versus host disease; bone marrow transplant; leukaemia; vaccine;

KW diagnosis; aplastic anaemia; immune deficiency disease.

XX Homo sapiens.

XX Key Location/Qualifiers

XX MISC-difference 3 /label="His, Arg"

XX WO9905174-A1.

XX 04-FEB-1999.

XX 23-JUL-1998; 98WO-NL00425.

XX 23-JUL-1997; 97EP-0202303.

XX (UYLE-) RIJXSUNIV LEIDEN.

XX Engelhard VH, Goulmy EAJM, Hunt DF;

XX WPI, 1999-153312/13.

XX A new minor histocompatibility antigen, HA-1 - useful to treat immune diseases and prevent rejection and host versus graft disease in bone marrow and organ transplantation

XX Claim 1; Page 32; 47pp; English.

XX The present sequence represents a new peptide (P1) constituting a T-cell epitope obtainable from the minor histocompatibility antigen HA-1. The peptide is immunogenic and can be used as part of a vaccine. P1 is used as a medicine, to induce tolerance for transplants, prevent rejection and/or graft versus host disease, or to treat (auto) immune diseases. In particular it can be used with bone marrow transplantation, in the treatment of severe aplastic anaemia, leukaemia, and immune deficiency diseases.

XX Sequence 9 AA:

Query Match 94.7%; Score 36; DB 20; Length 9;  
Best Local Similarity 100.0%; Pred. No. 7.8e+05;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VLXDDLLEA 9  
1 VLHDDLLEA 9

RESULT 5  
AAW99196  
ID AAW99196 standard; peptide; 9 AA.

XX

```

XX PR 23-JUL-1997; 97EP-0202303.
XX XX
XX PA (UYLE-) RIJKSUNIV LEIDEN.
XX PI Engelhard VH, Goulmv EAJM, Hunt DF;
XX DR WPI; 1999-153312/13.
XX XX
XX PT A new minor histocompatibility antigen, HA-1 - useful to treat
XX PT immune diseases and prevent rejection and host versus graft disease
XX PT in bone marrow and organ transplantation
XX PT
XX PS Disclosure; Page 15; 47pp: English.
XX XX
XX CC The present sequence represents a new peptide (PI) constituting a T-cell
XX CC epitope obtainable from the minor histocompatibility antigen HA-1. The
XX CC peptide is immunogenic and can be used as part of a vaccine. PI is used
XX CC as a medicine, to induce tolerance for transplants, prevent rejection
XX CC and/or graft versus host disease, or to treat (auto) immune diseases.
XX CC In particular it can be used with bone marrow transplantation, in the
XX CC treatment of severe aplastic anaemia, leukaemia, and immune deficiency
XX CC diseases.
XX
SQ Sequence 9 AA:
XX
XX Query Match 94.7%; Score 36; DB 20; Length 9;
XX Best Local Similarity 88.9%; Pred. No. 7.8e+05;
XX Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX Oy 1 VLKDXLEA 9
XX |||||||
XX 1 VLKDXLEA 9
XX
XX RESULT 7
XX AAW9198
XX ID AAW9198 standard: peptide; 13 AA.
XX AC AAW9198;
XX XX
XX DT 20-MAY-1999 (first entry)
XX XX
XX DE DH cell KIA0223 protein sequence.
XX XX
XX KW Minor histocompatibility antigen, HA-1, T-cell epitope; immunological;
XX KW graft versus host disease; bone marrow transplant; leukaemia; vaccine;
XX KW diagnosis; aplastic anaemia; immune deficiency disease.
XX XX
XX OS Synthetic.
XX XX
XX PN WO905174-A1.
XX XX
XX PD 04-FEB-1999.
XX XX
XX PF 23-JUL-1998; 98WO-NL00425.
XX XX
XX PR 23-JUL-1997; 97EP-0202303.
XX XX
XX PA (UYLE-) RIJKSUNIV LEIDEN.
XX XX
XX PI Engelhard VH, Goulmv EAJM, Hunt DF;
XX XX
XX DR WPI; 1999-153312/13.
XX DR N-PSDB; AAX19408.
XX XX
XX PT A new minor histocompatibility antigen, HA-1 - useful to treat
XX PT immune diseases and prevent rejection and host versus graft disease
XX PT in bone marrow and organ transplantation
XX PT
XX PS Disclosure; Page 31; 47pp: English.
XX XX
XX CC The present invention describes a new peptide (PI) constituting a T-cell

```

CC epitope obtainable from the minor histocompatibility antigen HA-1. The  
 CC peptide is immunogenic and can be used as part of a vaccine. PI is used  
 CC as a medicine, to induce tolerance for transplants, prevent rejection  
 CC and/or graft versus host disease, or to treat (auto) immune diseases.  
 CC In particular it can be used with bone marrow transplantation, in the  
 CC treatment of severe aplastic anaemia, leukaemia, and immune deficiency  
 CC diseases. The present sequence represents a KIAA0223 sequence given  
 CC in the present invention.

XX Sequence 13 AA:

Query Match 94.7%; Score 36; DB 20; Length 13;  
 Best Local Similarity 88.9%; Pred. No. 0.69;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLXDDLLEA 9  
 || |||||  
 Db 3 VLRDDILLEA 11

# RESULT 8

AAW9199 standard; peptide; 13 AA.

XX AAW9199;

XX 20-MAY-1999 (first entry)

DE VR cell KIAA0223 protein sequence.

XX Minor histocompatibility antigen; HA-1; T-cell epitope; immunological;  
 KW graft versus host disease; bone marrow transplant; leukaemia; vaccine;  
 KW diagnosis; aplastic anaemia; immune deficiency disease.

XX Synthetic.

XX WO9905174-A1.

XX 04-FEB-1999.

XX 23-JUL-1998; 98WO-NL00425.

XX 23-JUL-1997; 97EP-0202303.

XX (UYLE-) RIJXSUNIV LEIDEN.

XX Engelhard VH, Goulmey EAJM, Hunt DF;

XX WPI: 1999-153312/13.

XX N-PSDB; AAX19409.

XX A new minor histocompatibility antigen, HA-1 - useful to treat  
 PT immune diseases and prevent rejection and host versus graft disease  
 PT in bone marrow and organ transplantation

XX Disclosure; Page 31; 47pp; English.

XX The present invention describes a new peptide (PI) constituting a T-cell  
 CC epitope obtainable from the minor histocompatibility antigen HA-1. The  
 CC peptide is immunogenic and can be used as part of a vaccine. PI is used  
 CC as a medicine, to induce tolerance for transplants, prevent rejection  
 CC and/or graft versus host disease, or to treat (auto) immune diseases.  
 CC In particular it can be used with bone marrow transplantation, in the  
 CC treatment of severe aplastic anaemia, leukaemia, and immune deficiency  
 CC diseases. The present sequence represents a KIAA0223 sequence given  
 CC in the present invention.

XX Sequence 13 AA:

Query Match 94.7%; Score 36; DB 20; Length 13;  
 Best Local Similarity 88.9%; Pred. No. 0.69;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLXDDLLEA 9  
 || |||||  
 Db 3 VLRDDILLEA 11

# RESULT 9

AAW97414 standard; protein; 13 AA.

XX AAW97414;

XX 20-MAY-1999 (first entry)

DE KIAA0223 polymorphism in HA-1 negative homozygous individuals.

XX T-cell epitope; minor histocompatibility antigen HA-1; vaccine;  
 KW transplant rejection; Graft-versus-Host Disease; autoimmune disease;  
 KW neoplastic haematopoietic cell; KIAA0223 polymorphism.

XX Homo sapiens.

XX WO9905173-A1.

XX 04-FEB-1999.

XX 23-JUL-1998; 98WO-NL00424.

XX 23-JUL-1997; 97EP-0202303.

XX (UYLE-) RIJXSUNIV LEIDEN.

XX Engelhard VH, Goulmey EAJM, Hunt DF;

XX WPI: 1999-142855/12.

XX N-PSDB; AAX16080.

XX Immunogenic peptide from minor histocompatibility antigen HA-1 -  
 PT useful for inducing tolerance to transplants and prevent rejection  
 PT or graft-versus-host disease

XX Disclosure; Page 38; 57pp; English.

XX The present sequence represents the KIAA0223 polymorphism in HA-1  
 CC negative homozygous individuals. The specification describes an  
 CC immunogenic peptide constituting a T-cell epitope, obtainable from the  
 CC minor histocompatibility antigen HA-1. The peptide can be used in  
 CC vaccines or pharmaceutical formulations as medicines to induce  
 CC tolerance for transplants so as to prevent rejection and/or  
 CC Graft-versus-Host Disease, or to treat autoimmune diseases. Neoplastic  
 CC haematopoietic cells presenting the peptides, in an HLA class I context,  
 CC can be eliminated after specific recognition of the peptides. The  
 CC peptides can also be used to raise antibodies, T-cell receptor,  
 CC B- and T-cells.

XX Sequence 13 AA:

Query Match 94.7%; Score 36; DB 20; Length 13;  
 Best Local Similarity 88.9%; Pred. No. 0.69;  
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLXDDLLEA 9  
 || |||||  
 Db 3 VLRDDILLEA 11

# RESULT 10

AAW97415 standard; protein; 13 AA.

XX AAW97415;

XX 20-MAY-1999 (first entry)





PT vaccinating against and diagnosing infections, especially useful for  
PT treating acne vulgaris -

XX Example 1; SEQ ID No 14421; 1069pp; English.

XX Sequences AAU939105-AAU68017 represent Propionibacterium acnes immunogenic  
CC polypeptides. The proteins and their associated DNA sequences are used in  
CC the treatment, prevention and diagnosis of medical conditions caused by  
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,  
CC pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.  
CC P. acnes is also involved in infections of bone, joints and the central  
CC nervous system, however it is particularly involved in the inflammatory  
CC lesions associated with acne vulgaris. A method for detecting the  
CC presence or absence of P. acnes in a patient comprises contacting a  
CC sample with a binding agent that binds to the proteins of the invention  
CC and determining the amount of bound protein in the sample. The  
CC polypeptides may be used as antigens in the production of antibodies  
CC specific for P. acnes proteins. These antibodies can be used to  
CC downregulate expression and activity of P. acnes polypeptides and  
CC therefore treat P. acnes infections. The antibodies may also be used as  
CC diagnostic agents for determining P. acnes presence, for example, by  
CC enzyme linked immunosorbent assay (ELISA).  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX

XX Sequence 226 AA:

Query Match 86.8%; Score 33; DB 22; Length 226;

Best Local Similarity 87.5%; Pred. No. 55;  
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 LXDLEA 9  
| | | | |  
DB 20 LXDLEA 27

RESULT 13

AB90575 ID ABB90575 standard; Protein; 499 AA.

AC ABB90575;

DT 29-JUL-2002 (first entry)

XX Chlamydia pneumoniae cp6664 protein, SEQ ID NO:99.

XX Chlamydia pneumoniae; antigen; immunogen; vaccine; diagnosis;  
KW human respiratory disease; cardiovascular disease; atherosclerosis;  
KW coronary artery disease; carotid artery stenosis; myocardial infarction;  
KW cerebrovascular disease; aortic aneurysm; claudication; stroke;  
KW strain CWL029.

OS Chlamydia pneumoniae.

PN WO200202606-A2.

PD 10-JAN-2002.

PF 03-JUL-2001; 2001WO-IB01445.

XX 03-JUL-2000; 2000GB-0016363.

PR 11-JUL-2000; 2000GB-0017047.

PR 21-JUL-2000; 2000GB-0017983.

PR 07-AUG-2000; 2000GB-0019368.

PR 18-AUG-2000; 2000GB-0020440.

PR 14-SEP-2000; 2000GB-0022583.

PR 10-NOV-2000; 2000GB-0027549.

PR 22-DEC-2000; 2000GB-0031706.

XX (CHIR-) CHIRON SPA.

PA Ratil G, Grandi G;

XX PI

XX WPI: 2002-154726/20.  
DR N-PSDB: ABL91233.

XX Novel Chlamydia pneumoniae protein useful in the manufacture of a  
PT medicament for treatment or prevention of infection due to Chlamydia,  
PT preferably Chlamydia pneumoniae, and for diagnostic purposes -

PS Claim 1; Page 89-90; 364pp; English.

XX Sequences ABB90526-ABB90715 represent novel proteins from Chlamydia  
CC pneumoniae (strain CWL029), and ABL91184-ABL91373 represent DNA encoding  
CC them. The proteins are predicted to be immunogenic and may therefore be  
CC useful in vaccine production and for diagnostic purposes. Chlamydia  
CC pneumoniae is a common cause of respiratory disease in humans, and is  
CC also involved in the development of cardiovascular diseases such as  
CC atherosclerosis, coronary artery disease, carotid artery stenosis,  
CC myocardial infarction, cerebrovascular disease, aortic aneurysm,  
CC claudication and stroke. The proteins and nucleic acids of the invention  
CC may be used in vaccines and pharmaceutical compositions for the  
CC prevention or treatment of chlamydial infections, particularly Chlamydia  
CC pneumoniae infections. The proteins may also be used in the detection of  
CC Chlamydia pneumoniae, and the nucleic acids may be used in PCR, branched  
CC DNA probe assay or blotting techniques for determining Chlamydia  
CC pneumoniae gene expression. The present sequence represents a  
CC specifically claimed Chlamydia pneumoniae protein of the invention.  
XX

XX Sequence 499 AA:

Query Match 86.8%; Score 33; DB 23; Length 499;

Best Local Similarity 77.8%; Pred. No. 1.3e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLXDLEA 9  
| | | | |  
DB 395 VLXDLEA 403

RESULT 14

AA92707 ID AA92707 standard; Protein; 4150 AA.

AC AA92707;

DT 29-AUG-2000 (first entry)

XX S. antibioticus 8,8a-deoxyoleandolide synthase I encoded by oleAI.

XX oleandolide; polyketide synthase; oleAI; oleAII;

KW PKS; type I; 8,8a-deoxyoleandolide synthase; modular; ketosynthase;  
KW acyl-transferase; acyl carrier protein; inactivated; polyketide;  
KW macrolactone; antibiotic; motilide; erythromycin.

XX Streptococcus antibioticus.

XX Location/Qualifiers

XX Key 9..1034

XX Region /label= loading\_module

XX Domain /label= RS-Q-domain

XX /note= "inactivated ketosynthase domain"

XX /label= AT-domain

XX /note= "malonyl-specific acyl transferase domain"

XX Domain /label= ACP-domain

XX /note= "acyl carrier protein domain"

XX Region /label= Extender\_module\_1

XX Domain /label= KSI-domain

XX /note= "ketosynthase domain"

XX 1580..1926

FT /label= ATR1\_domain  
 FT /note= "acyl transferase domain"  
 FT 2162..2349  
 FT Domain  
 FT /label= KR1\_domain  
 FT /note= "ketoreductase domain"  
 FT 2440..2526  
 FT Domain  
 FT /label= ACP1\_domain  
 FT /note= "acyl carrier protein domain"  
 FT 2545..4069  
 FT Region  
 FT /label= Extender\_module\_2  
 FT 2545..2973  
 FT Domain  
 FT /label= KS2\_domain  
 FT /note= "ketosynthase domain"  
 FT 3075..3420  
 FT Domain  
 FT /label= AT2\_domain  
 FT /note= "acyl transferase domain"  
 FT 3700..3884  
 FT Domain  
 FT /label= KR2\_domain  
 FT /note= "ketoreductase domain"  
 FT 3979..4075  
 FT Domain  
 FT /label= ACP2\_domain  
 FT /note= "acyl carrier protein domain"  
 FT  
 PN WO200026349-A2.  
 PD 11-MAY-2000.  
 XX  
 XX 22-OCT-1999; 99WO-US24478.  
 PF  
 PR 29-OCT-1998; 98US-0106100.  
 PR 16-FEB-1999; 99US-0120254.  
 XX  
 PA (KOSA-) KOSAN BIOSCIENCES INC.  
 PI  
 PI Betlach MC, Shah SK, McDaniel R, Tang L;  
 XX  
 XX WPI: 2000-365602/31.  
 DR N-PDB; AAA09469.  
 XX  
 PT Recombinant DNA compound encoding oleandolide polyketide synthase for  
 PT synthesizing polyketides comprising a coding sequence for a domain of a  
 PT loading module or any one of extender modules  
 XX  
 PS Disclosure: Page 27-28; 86pp; English.  
 XX  
 CC The oleandolide polyketide synthase (PKS), also known as  
 CC 8,8a-deoxyoleandolide synthase, is encoded by three open reading frames  
 CC (ORF), designated oleA1, oleA2 and oleA3. The PKS is a type I  
 CC "modular" enzyme, where each ORF encodes 2 extender modules and  
 CC the first ORF also encodes the loading module. Each module is composed  
 CC of at least a ketosynthase (KS), acyl-transferase (AT) and an  
 CC acyl carrier protein (ACP) domain. The oleandolide PKS loading module  
 CC contains an inactivated KS, called KS-Q, where Q is the abbreviation for  
 CC glutamine, present instead of the active site cysteine required for  
 CC activity. The large multifunctional PKS enzymes catalyze the biosynthesis  
 CC of polyketide macrocycles through multistep pathways involving  
 CC decarboxylative condensations between acylthioesters followed by cycles  
 CC of varying beta-carbon processing activities. The macrocyclic product of  
 CC the PKS, 8,8a-deoxyoleandolide, is further modified by epoxidation and  
 CC glycosylation to yield oleandomycin, an antibacterial polyketide. The  
 CC invention concerns an isolated recombinant DNA compound, comprising a  
 CC coding sequence for a domain of loading module or any one of extender  
 CC modules 1-4 or 1-6, including an oleandolide PKS operably linked to a  
 CC promoter. Also discussed are recombinant oleandolide PKS in which the  
 CC module 1 KS domain is inactivated by deletion or other mutation. In  
 CC particular, the inactivation is mediated by a change in the KS domain  
 CC that renders it incapable of binding substrate (the KS1-Q mutation),  
 CC rendered by mutation in the codon for the active site cysteine. The  
 CC oleandolide PKS is useful for synthesizing polyketides, which are useful  
 CC as antibiotics and molluscs. Heterologous expression of oleandolide PKS  
 CC in host cells such as *Streptomyces coelicolor* and *S. lividans* is also  
 CC made possible. Unmodified oleandolide compounds can be provided to the  
 CC cultures of *Saccharopolyspora erythraea* and converted to the

CC corresponding derivatives of erythromycin A-D.  
 XX  
 SQ Sequence 4150 AA:  
 Query Match 84.2%; Score 32; DB 21; Length 4150;  
 Best Local Similarity 87.5%; Pred. No. 1.9e+03;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Oy 2 LXDPLEA 9  
 Db 1781 LRDPLEA 1788  
 RESULT 15  
 ID AAB25210  
 XX AAB25210 standard; Protein; 149 AA.  
 AC AAB25210;  
 XX  
 DT 27-NOV-2000 (first entry)  
 XX  
 DE Eucaalyptus grandis cell signalling involved protein SEQ ID NO:529.  
 XX  
 KW Eucaalyptus grandis; Pinus radiata; Monterey pine; plant; modification;  
 KW plant cell signalling; modulation; transgenic plant; pathogen; growth;  
 KW environmental change; development; cell proliferation; differentiation;  
 KW elongation; survival; disease resistance; nutrient metabolism.  
 XX  
 OS Eucaalyptus grandis.  
 XX  
 PN WO200042171-A1.  
 PD 20-JUL-2000.  
 XX  
 PF 11-JAN-2000; 2000WO-US00724.  
 PR  
 PR 12-JAN-1999; 99US-0228986.  
 PR 01-NOV-1999; 99US-0162866.  
 XX  
 PA (GENE-) GENESTIS RES & DEV CORP LTD.  
 PI  
 PI Strabala TV, Nieuwenhuizen NJ;  
 XX  
 XX WPI: 2000-476052/41.  
 DR  
 XX  
 PT Isolated polynucleotide encoding a polypeptide involved in cell  
 PT signalling used for generating transgenic plants with modified responses  
 PT to external signals -  
 XX  
 PS Claim 3; Page 235; 527pp; English.  
 XX  
 CC AAA92923 to AAA97936 and AAB25100 to AAB25570 represent polynucleotide  
 CC and protein sequences isolated from eucaalyptus (*Eucaalyptus grandis*) or  
 CC pine (*Pinus radiata* also known as Monterey pine). The protein sequences  
 CC are involved in cell signalling. The polynucleotide and protein  
 CC sequences can be used to modify the response of plant cells to external  
 CC signals e.g. environmental changes or pathogens during the growth and  
 CC development of a plant. They can be used to modify cell proliferation,  
 CC differentiation, elongation and survival, resistance to disease and  
 CC altered metabolism. Examples of modifications which can be produced are  
 CC delayed senescence and senescence of leaves and flowers e.g. to  
 CC delay senescence and prolong the life of cut flowers or enhance  
 CC senescence of reproductive organs to engineer sterile plants. Other  
 CC modifications can be used to delay senescence in selected cell types or  
 CC organs providing fruit and vegetables which have a longer shelf life  
 CC between harvest and consumption, or to decrease branching frequency in  
 CC forest tree species giving long stretches of valuable knot-free clear  
 CC wood which can be used in solid timber furniture and veneers.  
 XX  
 SQ Sequence 149 AA:  
 Query Match 81.6%; Score 31; DB 21; Length 149;  
 Best Local Similarity 66.7%; Pred. No. 91;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 VLDLLEA 9  
: | | | | |  
Db 73 ILDELLEA 81

## RESULT 16

ABR48245

ID ABR48245 standard; Protein: 276 AA.

AC ABR48245;

DT 05-FEB-2002 (first entry)

DE Listeria monocytogenes protein #949.

KW Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;

XX vitamin B12; bacterial infection; disease.

OS Listeria monocytogenes.

XX WO200177335-A2.

XX 18-OCT-2001.

XX 11-APR-2001; 2001WO-FR01118.

XX 11-APR-2000; 2000FR-0004629.

XX (INSP) INST PASTEUR.

PI Buchrieser C, Frangeul L, Couve E, Rusniok C, Psini H, Dehoux P;  
PI Dussurget O, Chelouani F, Nedjati H, Glaser P, Kunst F, Cossart P;  
PI Daniels J, Goebel W, Kreft J, Kuhn M, Ng E, Vazquez-Boland JA;  
PI Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A;  
PI Chakraborty T, Dommann E, Hain T, Berche P, Charbit A, Durant L;  
PI Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N;  
PI Madueno E, De Padlos B, Wenland J, Kaerst U, Entian K, Hauf J;  
PI Rose M, Voss H;

XX WPI: 2002-010914/01.

PT Genomic sequence for *Listeria monocytogenes*, useful e.g. for treatment  
PT and prevention of *Listeria* and related bacterial infections, and  
PT related polypeptides

PS Claim 6: SEQ ID No 950; 192pp; French.

XX The present invention relates to the genome sequence of *Listeria*  
CC monocytogenes EGD-e (see ABR48245). The genome sequence and fragments of  
CC it are useful for selecting probes and primers for detecting genes in *L.*  
CC monocytogenes and related organisms, and for studying genetic  
CC polymorphisms and other genomes. The present sequence is a protein  
CC encoded by the genome sequence of the present invention. Proteins  
CC expressed from the genome sequence are useful for raising specific  
CC antibodies, identification of *L. monocytogenes* and related organisms, and  
CC for biosynthesis and biodegradation, especially biosynthesis of vitamin  
CC B12. The genome sequence and proteins encoded by it are also useful for  
CC selecting compounds that regulate gene expression and cell replication  
CC and modulate *L. monocytogenes*-related diseases. In addition, the genome  
CC sequence and proteins encoded by it are useful in pharmaceutical and  
CC vaccines compositions for the treatment or prevention of infections by *L.*  
CC monocytogenes and related organisms.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.

SQ Sequence 276 AA;

Query Match 81.6%; Score 31; DB 23; Length 276;  
Best Local Similarity 75.0%; Pred. No. 1.7e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 LXDLLLEA 9  
: | | | | |  
Db 165 ISDDLLEA 172

## RESULT 17

AAU72987

ID AAU72987 standard; Protein: 289 AA.

AC AAU72987;

DT 12-MAR-2002 (first entry)

DE Neisseria meningitidis virulence protein #77.

XX Meningitis; virulence; gene; antibacterial; vaccine; veterinary;

KW infection; Gram-negative bacteria; antimicrobial.

XX Neisseria meningitidis.

XX WO200185772-A2.

XX 15-NOV-2001.

XX 08-MAY-2001; 2001WO-GB02003.

XX 08-MAY-2000; 2000GB-0011108.

XX (MICR-) MICROSCIENCE LTD.

XX Tang C;

XX WPI: 2002-066593/09.

XX N-PSDB; AAS97272.

PT New peptide encoded by operon including virulence genes of *Neisseria*  
PT meningitidis, useful as vaccine component for treating or preventing  
PT meningitis and for identifying antimicrobial drug

PS Claim 4: Page 319-320; 423pp; English.

XX The invention relates to a peptide (I) encoded by an operon (II) of  
CC *Neisseria meningitidis* including virulence genes, or a related molecule  
CC having a 40% sequence similarity at the peptide or nucleotide level in a  
CC Gram-negative bacterium, or its functional fragment, for therapeutic or  
CC diagnostic use. (I) and (II) are useful in the manufacture of a  
CC medicament for treating or preventing a condition (e.g., meningitis)  
CC associated with infection by *Neisseria* or Gram-negative bacteria. The  
CC product is useful for veterinary treatment and in a screening assay for  
CC the identification of an antimicrobial drug. The vaccines have  
CC prophylactic applications. AAU72911-AAU73014 represent N. meningitidis  
CC virulence proteins of the invention.

SQ Sequence 289 AA;

Query Match 81.6%; Score 31; DB 23; Length 289;  
Best Local Similarity 55.6%; Pred. No. 1.8e+02;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 VLDLLEA 9  
: | | | | |  
Db 130 ILADLLEA 138

## RESULT 18

AAU73000

ID AAU73000 standard; Protein: 289 AA.

AC AAU73000;

DT 12-MAR-2002 (first entry)

DE Neisseria meningitidis virulence protein #90.  
 XX  
 KW Meningitis: virulence; gene: antibacterial; vaccine; veterinary;  
 XX infection; Gram-negative bacteria; antimicrobial.  
 OS Neisseria meningitidis.  
 XX  
 PN WO200185772-A2.  
 XX  
 PD 15-NOV-2001.  
 XX  
 PF 08-MAY-2001; 2001WO-GB02003.  
 XX  
 PR 08-MAY-2000; 2000GB-0011108.  
 XX  
 PA (MICR-) MICROSCIENCE LTD.  
 XX  
 PI Tang C;  
 XX  
 DR WPI: 2002-066593/09.  
 DR N-PSDB: AAS97285.  
 XX  
 PT New peptide encoded by operon including virulence genes of Neisseria  
 PT meningitidis, useful as vaccine component for treating or preventing  
 PT meningitis and for identifying antimicrobial drug -  
 XX  
 PS Claim 4: Page 366-367; 423pp: English.  
 XX  
 CC The invention relates to a peptide (I) encoded by an operon (II) of  
 CC Neisseria meningitidis including virulence genes, or a related molecule  
 CC having a 40% sequence similarity at the peptide or nucleotide level in a  
 CC Gram-negative bacterium, or its functional fragment, for therapeutic or  
 CC diagnostic use. (I) and (II) are useful in the manufacture of a  
 CC associated with infection or preventing a condition (e.g., meningitis)  
 CC product is useful for veterinary treatment and in a screening assay for  
 CC the identification of an antimicrobial drug. The vaccines have  
 CC prophylactic applications. AAU72911-AAU73014 represent N. meningitidis  
 CC virulence proteins of the invention.  
 CC  
 SQ Sequence 289 AA;  
 XX  
 XX  
 Query Match 81.6%; Score 31; DB 23; Length 289;  
 Best Local Similarity 55.6%; Pred. No. 1.8e+02;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 VLXDDLLEA 9  
 : | | | | :  
 Db 130 ILADDLIDA 138  
 XX  
 RESULT 19  
 ABB59345  
 ID ABB59345 standard; Protein; 1042 AA.  
 XX  
 AC ABB59345;  
 XX  
 DT 26-MAR-2002 (first entry)  
 XX  
 DE Drosophila melanogaster polypeptide SHQ ID NO 4827.  
 XX  
 KW Drosophila; developmental biology; cell signalling; insecticide;  
 KW pharmaceutical.  
 XX  
 OS Drosophila melanogaster.  
 XX  
 PN WO200171042-A2.  
 XX  
 PD 27-SEP-2001.  
 XX  
 PF 23-MAR-2001; 2001WO-US09231.  
 XX  
 PR 23-MAR-2000; 2000US-191637P.  
 PR

PR 11-JUL-2000; 2000US-0614150.  
 XX  
 PA (PEKE ) PE CORP NY.  
 XX  
 PI Venter JC, Adams M, Li PWD, Myers EW;  
 XX  
 DR WPI: 2001-656860/75.  
 DR N-PSDB: ABL03448.  
 XX  
 PT New isolated nucleic acid detection reagent for detecting 1000 or more  
 PT genes from Drosophila and for elucidating cell signalling and cell-cell  
 PT interactions -  
 XX  
 PS Disclosure; SHQ ID NO 4827; 21pp + Sequence Listing; English.  
 XX  
 CC The invention relates to an isolated nucleic acid detection reagent  
 CC capable of detecting 1000 or more genes from Drosophila. The invention is  
 CC useful in developmental biology and in elucidating cell signalling and  
 CC cell-cell interactions in higher eukaryotes for the development of  
 CC insecticides, therapeutics and pharmaceutical drugs. The invention  
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
 CC sequences (ABL01840-ABL16175) and the encoded proteins  
 CC (AB57737-AB872072).  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 CC  
 SQ Sequence 1042 AA;  
 XX  
 XX  
 Query Match 81.6%; Score 31; DB 22; Length 1042;  
 Best Local Similarity 77.8%; Pred. No. 7e+02;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 VLXDDLLEA 9  
 : | | | | :  
 Db 595 VLDDALEA 603  
 XX  
 RESULT 20  
 AAU86478  
 ID AAU86478 standard; Protein; 94 AA.  
 XX  
 AC AAU86478;  
 XX  
 DT 21-MAY-2002 (first entry)  
 XX  
 DE Novel human connective tissue related polypeptide #44.  
 XX  
 KW Human; connective tissue related disorder; cancer; cytostatic.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200155343-A1.  
 XX  
 PD 02-AUG-2001.  
 XX  
 PF 17-JAN-2001; 2001WO-US01322.  
 XX  
 PR 31-JAN-2000; 2000US-0179065.  
 PR 04-FEB-2000; 2000US-0180628.  
 PR 24-FEB-2000; 2000US-0184664.  
 PR 02-MAR-2000; 2000US-0186350.  
 PR 16-MAR-2000; 2000US-0189874.  
 PR 17-MAR-2000; 2000US-0190076.  
 PR 18-APR-2000; 2000US-0198123.  
 PR 19-MAY-2000; 2000US-0205515.  
 PR 07-JUN-2000; 2000US-0209467.  
 PR 28-JUN-2000; 2000US-0214886.  
 PR 30-JUN-2000; 2000US-0215135.  
 PR 07-JUL-2000; 2000US-0216647.  
 PR 07-JUL-2000; 2000US-0216880.  
 PR 11-JUL-2000; 2000US-0217487.  
 PR 11-JUL-2000; 2000US-0217496.  
 PR



XX Sequence 94 AA;  
SQ

Query Match  
Best Local Similarity 78.9%; Score 30; DB 22; Length 94;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 LXDDLEA 9  
Db 21 LSDDLLES 28

RESULT 21

AAG73861  
ID AAG73861 standard; Protein; 328 AA.

XX AAG73861;

DT 03-SEP-2001 (first entry)

DE Human colon cancer antigen protein SEQ ID NO:4625.

KW Human; colon cancer; colon cancer antigen; diagnosis; detection;  
colorectal carcinoma; chromosome 12.

OS Homo sapiens.

PN WO200122920-A2.

PD 05-APR-2001.

PF 28-SEP-2000; 2000WO-US26524.

PR 29-SEP-1999; 99US-0157137.

PR 03-NOV-1999; 99US-0163280.

PA (HDNA-) HDNA GENOME SCI INC.

PI Ruben SM, Barash SC, Birse CE, Rosen CA;

DR WPI: 2001-235357/24.

DR N-PSDB: AAH33292.

XX Claim 11; Page 6423-6425; 9803pp; English.

XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon  
XX cancer-associated nucleic acid molecules (N) and proteins (P), where  
XX the proteins are collectively known as colon cancer antigens. The colon  
XX cancer antigens have cytostatic activity and can be used in gene  
XX therapy and vaccine production. N and P may be used in the prevention,  
XX diagnosis and treatment of diseases associated with inappropriate P  
XX expression. For example, N and P may be used to treat disorders  
XX associated with decreased expression by rectifying mutations or deletions  
XX in a patient's genome that affect the activity of P by expressing P.  
XX Inactive proteins or to supplement the activity of P by expressing P.  
XX Additionally, N may be used to produce the colon cancer-associated P,  
XX by inserting the nucleic acids into a host cell and culturing the cell  
XX to express the proteins. N and P can be used in the prevention, diagnosis  
XX and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204  
XX and AAB77789 represent sequences used in the exemplification of the  
XX present invention.  
XX N.B. Pages 666 to 682 and page 7053 of the sequence listing were  
XX missing at time of publication, meaning no sequences are present for  
XX SEQ ID NO:1027 to 1052, 7921 and 7922.

SQ Sequence 328 AA;

Query Match  
Best Local Similarity 78.9%; Score 30; DB 22; Length 328;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 VLXDDLE 8  
Db 183 VLKDDLE 190

RESULT 22

AAB96742  
ID AAB96742 standard; Protein; 623 AA.

XX AAB96742;

DT 29-OCT-2001 (first entry)

DE Putative P. abyssi type I site specific endonuclease, subunit M.

KW Hyperthermophilic archaeon; hyperthermophilic protein.

OS Pyrococcus abyssi.

PN FR2792651-A1.

PD 27-OCT-2000.

PF 21-APR-1999; 99FR-0005034.

PR 21-APR-1999; 99FR-0005034.

PA (CNRS) CNRS CENT NAT RECH SCI.

PI (IFRE-) IFREMER INST FR RECH EXPL MER.

PI Forterre P, Thiery JC, Prieur D, Dietrich J, Lecompte O;

PI Querellou J, Weissenbach J, Saurin W, Hellig R;

DR WPI: 2001-126236/14.

XX New nucleotide sequences isolated from Pyrococcus abyssi encode  
XX proteins useful in industry -

XX Claim 7; Pages 1507-1509; 1657pp; French.

XX The present invention relates to the genomic sequence of Pyrococcus  
XX abyssi (see AAF66431 and AAH41223-7) and P. abyssi proteins. P. abyssi is  
XX a hyperthermophilic archaeon, which is isolated from deep sea  
XX hydrothermal vents. The present sequence is one such P. abyssi protein.  
XX The proteins of the present invention have various potential industrial  
XX uses, since the proteins are stable at very high temperatures, some up to  
XX 110 degrees centigrade.  
XX Note: This patent is in the same patent family as WO200065062, which  
XX contains additional sequences as shown in AAB99132-AAB99143,  
XX AAH75903-AAH75920 and AAG66436.

SQ Sequence 623 AA;

Query Match  
Best Local Similarity 78.9%; Score 30; DB 22; Length 623;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 VLXDDLEA 9  
Db 472 IVEDDLEA 480

RESULT 23

AAAY58574  
ID AAAY58574 standard; Protein; 1410 AA.

XX AAAY58574;

DT 10-APR-2000 (first entry)

DE Sorangium cellulosum non-ribosomal peptide synthetase EPOS P.



XX EPOS P, non-ribosomal peptide synthetase; polyketide synthase;  
 KW epoethione biosynthesis; thiazole ring formation; taxol substitute;  
 XX anticancer.  
 XX Sorangium cellulosum.  
 OS WO966028-A2.  
 PN  
 XX  
 PD 23-DEC-1999.  
 PF 16-JUN-1999; 99WO-EP04171.  
 XX  
 PR 18-JUN-1998; 980S-0099504.  
 XX 24-SEP-1998; 98US-0101631.  
 PR 05-FEB-1999; 99US-0118906.  
 XX  
 PA (NOVS ) NOVARTIS AG.  
 PI (NOVS ) NOVARTIS-ERFINDUNGEN VERW GES MBH.  
 PI Schupp T, Ligon JM, Molnar I, Zirkle R, Goerlach J, Cyr D;  
 DR WPI; 2000-097741/08.  
 DR N-PSDB; AA255887.  
 XX  
 PT New isolated epoethione synthase genes, used for the recombinant  
 PT production of epoethione for use in cancer therapy -  
 PS  
 PS Claim 12; Page 109-113; 174pp; English.  
 XX  
 CC This sequence represents a Sorangium cellulosum non-ribosomal peptide  
 CC synthetase EPOS P, which is one of several epoethione biosynthetic  
 CC enzymes encoded by a 68.75 kb contig. Epoethiones A and B are  
 CC 16-membered macrocyclic polyketides with an acylcysteine-derived  
 CC starter unit; polyketides being synthesised from two-carbon building  
 CC blocks, the beta-carbon of which always carries a keto group. Each round  
 CC of two-carbon addition is carried out by a complex of enzymes known as  
 CC the polyketide synthase in a manner similar to fatty acid biosynthesis.  
 CC EPOS A (AAV58573) and EPOS P (AAV58574) are involved in formation of  
 CC the thiazole ring formation of epoethiones, and EPOS B, EPOS C, EPOS D  
 CC and EPOS E (AAV58575-V58578) are involved in polyketide backbone  
 CC formation. EPO F (AAV58582) and Orf14 (AAV58593) are thought to be  
 CC the proteins Orf 3 (AAV58582) and Orf14 (AAV58593) are thought to be  
 CC involved in transport. Epoethiones mimic the biological activity of  
 CC taxol, and may be substituted for taxol in cancer chemotherapeutic  
 CC compositions. Epoethiones exhibit a much lower drop in potency against a  
 CC multiply drug-resistant cell line compared with taxol, and are  
 CC considerably less efficiently exported from such cells by the multidrug  
 CC resistance protein (MDR, or P-glycoprotein). Despite the potential of a  
 CC epoethiones as anticancer agents, they are problematical to produce on a  
 CC large scale. Epoethiones are too complex for industrial scale chemical  
 CC synthesis, and Sorangium cellulosum is difficult to ferment, producing  
 CC poor yields of epoethiones. The nucleic acids of the invention may be  
 CC used for the recombinant production of epoethiones in a heterologous host  
 CC that is more amenable to fermentation.  
 CC  
 SO Sequence 1410 AA;  
 QY  
 DB 474 LDDMLEA 481  
 OY 2 LDDMLEA 9  
 ID 1111111  
 AC AAU02010 standard; Protein; 75 AA.  
 AC AAU02010;  
 XX  
 DT 29-AUG-2001 (first entry)

XX Gene #37 human secreted protein homologous amino acid sequence.  
 DE  
 XX  
 KW Human secreted protein; autoimmune disorder; hyperproliferative disorder;  
 KW cardiovascular disorder; cerebrovascular disorder; angiogenesis;  
 KW nervous system disorder; bacterial infection; viral infection;  
 KW fungal infection; ocular disorder; wound healing; tissue regeneration;  
 KW epithelial cell proliferation; skin ageing; chemotaxis; IgG Fc region.  
 XX  
 OS Homo sapiens.  
 PN  
 XX  
 PD WO200123598-A1.  
 PD 05-APR-2001.  
 XX  
 PF 26-SEP-2000; 2000WO-US26324.  
 XX  
 PR 27-SEP-1999; 99US-0155807.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 PI Komatsoulis G, Ruben SM, Rosen CA;  
 DR WPI; 2001-281684/29.  
 XX  
 PT Forty one nucleic acid molecules encoding human secreted proteins, useful  
 PT in the prevention, treatment and diagnosis of cancer, immune disorders,  
 PT cardiovascular disorders and neurological diseases -  
 PS  
 PS Disclosure; Page 511-512; 518pp; English.  
 XX  
 CC The sequence represents a polypeptide homologous to a human secreted  
 CC protein of the invention. Secreted proteins and their related nucleic  
 CC acids can be used in the diagnosis of or susceptibility to a pathological  
 CC condition by determining the presence or absence of a mutation in a  
 CC nucleic acid or the presence or amount of expression of a secreted  
 CC protein. The sequences are used to prevent, treat or ameliorate a medical  
 CC condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs,  
 CC chickens or sheep. The antibodies to the polypeptides can also be used in  
 CC assays (ELISA). The disorders include autoimmune diseases e.g. rheumatoid  
 CC arthritis, hyperproliferative disorders e.g. cardiac arrest, cerebrovascular  
 CC disorders e.g. cerebral ischaemia, angiogenesis, nervous system disorders  
 CC e.g. Alzheimer's disease, infections caused by bacteria, viruses and  
 CC fungi and ocular disorders e.g. corneal infection. The peptides can also  
 CC be used to aid wound healing and epithelial cell proliferation, to help  
 CC prevent skin ageing due to sunburn, to maintain organs before  
 CC transplantation, to regenerate tissues, in chemotaxis and as a food  
 CC additive or preservative to alter storage capabilities.  
 CC  
 SO Sequence 75 AA;  
 QY  
 DB 28 DDLLEA 33  
 OY 4 DDLLEA 9  
 ID 1111111  
 AC AAG74726  
 AC AAG74726 standard; Protein; 97 AA.  
 AC AAG74726;  
 XX  
 DT 03-SEP-2001 (first entry)  
 XX  
 DE Human colon cancer antigen protein SEQ ID NO:5490.  
 XX

KW Human; colon cancer; colon cancer antigen; diagnosis; detection;  
 KW colorectal carcinoma.  
 OS Homo sapiens.  
 XX  
 PN WO200122920-A2.  
 XX  
 PD 05-APR-2001.  
 XX  
 PE 28-SEP-2000; 2000MO-US26524.  
 XX  
 PR 29-SEP-1999; 99US-0157137.  
 PR 03-NOV-1999; 99US-0163280.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Ruben SM, Barash SC, Birse CE, Rosen CA;  
 DR MPI; 2001-235357/24.  
 DR N-PSDB; AAH34131.  
 XX  
 PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,  
 PT useful for preventing, diagnosing and/or treating colorectal cancers -  
 XX  
 PS Claim 11; Page 7086; 9803pp; English.  
 XX  
 CC AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon  
 CC cancer-associated nucleic acid molecules (N) and proteins (P), where  
 CC the proteins are collectively known as colon cancer antigens. The colon  
 CC cancer antigens have cytostatic activity and can be used in gene  
 CC therapy and vaccine production. N and P may be used in the prevention,  
 CC diagnosis and treatment of diseases associated with inappropriate P  
 CC expression. For example, N and P may be used to treat disorders  
 CC associated with decreased expression by rectifying mutations or deletions  
 CC in a patient's genome that affect the activity of P by expressing  
 CC inactive proteins or to supplement the patients own production of P.  
 CC Additionally, N may be used to produce the colon cancer-associated Ps,  
 CC by inserting the nucleic acids into a host cell and culturing the cell  
 CC to express the proteins. N and P can be used in the prevention, diagnosis  
 CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204  
 CC present invention  
 CC N.B. Pages 666 to 682 and page 7053 of the sequence listing were  
 CC missing at time of publication, meaning no sequences are present for  
 CC SEQ ID NO:1027 to 1052, 7921 and 7922.  
 XX  
 SQ Sequence 97 AA;  
 Query Match 76.3%; Score 29; DB 22; Length 97;  
 Best Local Similarity 75.0%; Pred. No. 1.5e+02;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 2 LXDDLEA 9  
 Db 59 LCDDLLEA 66  
 RESULT 26  
 AAG05155  
 ID AAG05155 standard; Protein: 144 AA.  
 AC AAG05155;  
 XX  
 DT 17-OCT-2000 (first entry)  
 XX  
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 1454.  
 XX  
 KW Protein identification; signal transduction pathway; metabolic pathway;  
 KW hybridisation assay; genetic mapping; gene expression control; promoter;  
 XX termination sequence.  
 OS Arabidopsis thaliana.  
 XX

PN EPI033405-A2.  
 XX  
 PD 06-SEP-2000.  
 XX  
 PE 25-FEB-2000; 2000EP-0301439.  
 XX  
 PR 25-FEB-1999; 99US-0121825.  
 PR 05-MAR-1999; 99US-0123180.  
 PR 09-MAR-1999; 99US-0123548.  
 PR 23-MAR-1999; 99US-0125788.  
 PR 25-MAR-1999; 99US-0126264.  
 PR 29-MAR-1999; 99US-0126785.  
 PR 01-APR-1999; 99US-0127462.  
 PR 06-APR-1999; 99US-0128234.  
 PR 08-APR-1999; 99US-0128714.  
 PR 16-APR-1999; 99US-0129845.  
 PR 19-APR-1999; 99US-0130077.  
 PR 21-APR-1999; 99US-0130449.  
 PR 23-APR-1999; 99US-0130510.  
 PR 28-APR-1999; 99US-0130891.  
 PR 30-APR-1999; 99US-0131449.  
 PR 30-APR-1999; 99US-0132048.  
 PR 04-MAY-1999; 99US-0132407.  
 PR 05-MAY-1999; 99US-0132484.  
 PR 06-MAY-1999; 99US-0132485.  
 PR 07-MAY-1999; 99US-0132486.  
 PR 11-MAY-1999; 99US-0132487.  
 PR 14-MAY-1999; 99US-0132863.  
 PR 14-MAY-1999; 99US-0134256.  
 PR 14-MAY-1999; 99US-0134218.  
 PR 14-MAY-1999; 99US-0134219.  
 PR 14-MAY-1999; 99US-0134321.  
 PR 18-MAY-1999; 99US-0134370.  
 PR 19-MAY-1999; 99US-0134768.  
 PR 20-MAY-1999; 99US-0134941.  
 PR 21-MAY-1999; 99US-0135124.  
 PR 24-MAY-1999; 99US-0135353.  
 PR 25-MAY-1999; 99US-0135629.  
 PR 27-MAY-1999; 99US-0136021.  
 PR 28-MAY-1999; 99US-0136392.  
 PR 01-JUN-1999; 99US-0136782.  
 PR 03-JUN-1999; 99US-0137222.  
 PR 04-JUN-1999; 99US-0137528.  
 PR 07-JUN-1999; 99US-0137502.  
 PR 08-JUN-1999; 99US-0137724.  
 PR 10-JUN-1999; 99US-0138094.  
 PR 10-JUN-1999; 99US-0138540.  
 PR 14-JUN-1999; 99US-0138847.  
 PR 16-JUN-1999; 99US-0139119.  
 PR 16-JUN-1999; 99US-0139452.  
 PR 17-JUN-1999; 99US-0139453.  
 PR 18-JUN-1999; 99US-0139454.  
 PR 18-JUN-1999; 99US-0139455.  
 PR 18-JUN-1999; 99US-0139456.  
 PR 18-JUN-1999; 99US-0139457.  
 PR 18-JUN-1999; 99US-0139458.  
 PR 18-JUN-1999; 99US-0139459.  
 PR 18-JUN-1999; 99US-0139460.  
 PR 18-JUN-1999; 99US-0139461.  
 PR 18-JUN-1999; 99US-0139462.  
 PR 18-JUN-1999; 99US-0139463.  
 PR 18-JUN-1999; 99US-0139750.  
 PR 21-JUN-1999; 99US-0139763.  
 PR 22-JUN-1999; 99US-0139817.  
 PR 23-JUN-1999; 99US-0139899.  
 PR 23-JUN-1999; 99US-0140353.  
 PR 24-JUN-1999; 99US-0140354.  
 PR 28-JUN-1999; 99US-0140695.  
 PR 29-JUN-1999; 99US-0140823.  
 PR 30-JUN-1999; 99US-0140991.  
 PR 01-JUL-1999; 99US-0141287.  
 PR 01-JUL-1999; 99US-0141842.



PR 05-MAR-1999; 99US-0123180.  
 PR 09-MAR-1999; 99US-0123548.  
 PR 23-MAR-1999; 99US-0125788.  
 PR 25-MAR-1999; 99US-0126264.  
 PR 29-MAR-1999; 99US-0126785.  
 PR 01-APR-1999; 99US-0127462.  
 PR 06-APR-1999; 99US-0128234.  
 PR 08-APR-1999; 99US-0128714.  
 PR 16-APR-1999; 99US-0129845.  
 PR 19-APR-1999; 99US-0130077.  
 PR 21-APR-1999; 99US-0130449.  
 PR 23-APR-1999; 99US-0130510.  
 PR 28-APR-1999; 99US-0130891.  
 PR 30-APR-1999; 99US-0131449.  
 PR 04-MAY-1999; 99US-0132048.  
 PR 05-MAY-1999; 99US-0132407.  
 PR 06-MAY-1999; 99US-0132484.  
 PR 06-MAY-1999; 99US-0132485.  
 PR 07-MAY-1999; 99US-0132486.  
 PR 11-MAY-1999; 99US-0132487.  
 PR 14-MAY-1999; 99US-0132863.  
 PR 14-MAY-1999; 99US-0134218.  
 PR 14-MAY-1999; 99US-0134219.  
 PR 14-MAY-1999; 99US-0134221.  
 PR 18-MAY-1999; 99US-0134370.  
 PR 19-MAY-1999; 99US-0134768.  
 PR 20-MAY-1999; 99US-0134941.  
 PR 21-MAY-1999; 99US-0135124.  
 PR 24-MAY-1999; 99US-0135353.  
 PR 25-MAY-1999; 99US-0135629.  
 PR 27-MAY-1999; 99US-0136021.  
 PR 28-MAY-1999; 99US-0136392.  
 PR 01-JUN-1999; 99US-0136782.  
 PR 03-JUN-1999; 99US-0137222.  
 PR 04-JUN-1999; 99US-0137528.  
 PR 07-JUN-1999; 99US-0137502.  
 PR 08-JUN-1999; 99US-0137724.  
 PR 10-JUN-1999; 99US-0138094.  
 PR 14-JUN-1999; 99US-0138540.  
 PR 14-JUN-1999; 99US-0138847.  
 PR 16-JUN-1999; 99US-0139119.  
 PR 16-JUN-1999; 99US-0139452.  
 PR 17-JUN-1999; 99US-0139453.  
 PR 18-JUN-1999; 99US-0139459.  
 PR 18-JUN-1999; 99US-0139457.  
 PR 18-JUN-1999; 99US-0139458.  
 PR 18-JUN-1999; 99US-0139459.  
 PR 18-JUN-1999; 99US-0139460.  
 PR 18-JUN-1999; 99US-0139461.  
 PR 18-JUN-1999; 99US-0139462.  
 PR 18-JUN-1999; 99US-0139463.  
 PR 18-JUN-1999; 99US-0139750.  
 PR 21-JUN-1999; 99US-0139763.  
 PR 22-JUN-1999; 99US-0139817.  
 PR 23-JUN-1999; 99US-0139819.  
 PR 23-JUN-1999; 99US-0140353.  
 PR 24-JUN-1999; 99US-0140354.  
 PR 28-JUN-1999; 99US-0140695.  
 PR 28-JUN-1999; 99US-0140823.  
 PR 29-JUN-1999; 99US-0140991.  
 PR 30-JUN-1999; 99US-0141287.  
 PR 01-JUL-1999; 99US-0141842.  
 PR 02-JUL-1999; 99US-0142154.  
 PR 06-JUL-1999; 99US-0142055.  
 PR 08-JUL-1999; 99US-0142390.  
 PR 09-JUL-1999; 99US-0142803.  
 PR 12-JUL-1999; 99US-0142920.  
 PR 13-JUL-1999; 99US-0143542.

PR 14-JUL-1999; 99US-0143624.  
 PR 15-JUL-1999; 99US-0144005.  
 PR 16-JUL-1999; 99US-0144085.  
 PR 16-JUL-1999; 99US-0144086.  
 PR 19-JUL-1999; 99US-0144325.  
 PR 19-JUL-1999; 99US-0144331.  
 PR 19-JUL-1999; 99US-0144332.  
 PR 19-JUL-1999; 99US-0144333.  
 PR 19-JUL-1999; 99US-0144334.  
 PR 19-JUL-1999; 99US-0144335.  
 PR 20-JUL-1999; 99US-0144352.  
 PR 20-JUL-1999; 99US-0144632.  
 PR 21-JUL-1999; 99US-0144884.  
 PR 21-JUL-1999; 99US-0144814.  
 PR 21-JUL-1999; 99US-0145086.  
 PR 22-JUL-1999; 99US-0145087.  
 PR 22-JUL-1999; 99US-0145088.  
 PR 22-JUL-1999; 99US-0145089.  
 PR 22-JUL-1999; 99US-0145192.  
 PR 23-JUL-1999; 99US-0145145.  
 PR 23-JUL-1999; 99US-0145218.  
 PR 26-JUL-1999; 99US-0145224.  
 PR 27-JUL-1999; 99US-0145276.  
 PR 27-JUL-1999; 99US-0145913.  
 PR 27-JUL-1999; 99US-0145918.  
 PR 28-JUL-1999; 99US-0145919.  
 PR 02-AUG-1999; 99US-0145951.  
 PR 02-AUG-1999; 99US-0146386.  
 PR 02-AUG-1999; 99US-0146388.  
 PR 03-AUG-1999; 99US-0146389.  
 PR 04-AUG-1999; 99US-0147038.  
 PR 05-AUG-1999; 99US-0147204.  
 PR 05-AUG-1999; 99US-0147302.  
 PR 05-AUG-1999; 99US-0147312.  
 PR 06-AUG-1999; 99US-0147260.  
 PR 06-AUG-1999; 99US-0147260.  
 PR 09-AUG-1999; 99US-0147303.  
 PR 09-AUG-1999; 99US-0147416.  
 PR 09-AUG-1999; 99US-0147493.  
 PR 10-AUG-1999; 99US-0148171.  
 PR 11-AUG-1999; 99US-0148171.  
 PR 12-AUG-1999; 99US-0148319.  
 PR 13-AUG-1999; 99US-0148341.  
 PR 13-AUG-1999; 99US-0148565.  
 PR 16-AUG-1999; 99US-0148684.  
 PR 17-AUG-1999; 99US-0149368.  
 PR 18-AUG-1999; 99US-0149175.  
 PR 20-AUG-1999; 99US-0149426.  
 PR 20-AUG-1999; 99US-0149722.  
 PR 20-AUG-1999; 99US-0149723.  
 PR 23-AUG-1999; 99US-0149929.  
 PR 23-AUG-1999; 99US-0149930.  
 PR 23-AUG-1999; 99US-0150566.  
 PR 26-AUG-1999; 99US-0150884.  
 PR 27-AUG-1999; 99US-0151065.  
 PR 27-AUG-1999; 99US-0151066.  
 PR 30-AUG-1999; 99US-0151080.  
 PR 31-AUG-1999; 99US-0151303.  
 PR 01-SEP-1999; 99US-0151338.  
 PR 07-SEP-1999; 99US-0151930.  
 PR 10-SEP-1999; 99US-0152363.  
 PR 13-SEP-1999; 99US-0153070.  
 PR 15-SEP-1999; 99US-0153758.  
 PR 16-SEP-1999; 99US-0154018.  
 PR 20-SEP-1999; 99US-0154039.  
 PR 22-SEP-1999; 99US-0154779.  
 PR 23-SEP-1999; 99US-0155139.  
 PR 24-SEP-1999; 99US-0155486.  
 PR 28-SEP-1999; 99US-0155659.  
 PR 29-SEP-1999; 99US-0156458.  
 PR 04-OCT-1999; 99US-0157117.

```

PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159325.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

```

Query Match 76.3%; Score 29; DB 21; Length 153;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 4 DDLLEA 9
DB 22 DDLLEA 27

```

```

RESULT 28
ABB97149
ID ABB97149 standard; Protein: 216 AA.
AC ABB97149;
XX 21-JUN-2002 (first entry)
DT
XX
XX Human tumour antigen related protein SEQ ID NO 51.
DE
XX
XX Human; tumour; antigen; HLA-A2; cytotoxic T cell; cytostatic; cancer;
KW vaccine.
XX
XX Homo sapiens.
OS
XX
XX WO200210369-A1.
PN
XX
XX 07-FEB-2002.
PD
XX
XX 30-JUL-2001; 2001WO-JP06526.
PF
XX
XX 31-JUL-2000; 2000JP-0231814.
PR
XX
XX (ITOH/) ITOH K.
PA
XX
XX Itoh K;
PI
XX
XX WPI; 2002-291857/33.
DR
XX
XX N-PSDB; ABL56077.
XX

```

```

PT Tumor antigens inducing and/or activating HLA-A2-restricted
PT tumor-specific cytotoxic T cells; useful in diagnosis of and screening
PT drugs e.g. cancer vaccines for specific treatment of pancreatic cancer
PT
PS
XX Claim 2; Page 107-108; 127pp; Japanese.
XX
CC The invention relates to a peptide comprising an amino acid sequence
CC selected from 44 fully defined amino acid sequences (ABB96906-ABB969549)
CC and a polypeptide comprising an amino acid sequence selected from the 9
CC fully defined amino acid sequences (ABB97143-ABB97151). The above
CC comprise a tumour antigen inducing or activating HLA-A2 restricted
CC tumour-specific cytotoxic T cells, which recognise HLA-A2 and a tumour
CC antigen peptide and is thus activated. The peptides and polypeptides
CC have cytostatic activity. The tumour antigen is useful in diagnosis of
CC and screening drugs for specific treatment of pancreatic cancer, colon
CC cancer and stomach cancer including in the form of vaccines. The present
CC sequence is that of a tumour antigen protein, useful to the invention.
CC
XX
XX Sequence 216 AA:

```

```

Query Match 76.3%; Score 29; DB 23; Length 216;
Best Local Similarity 62.5%; Pred. No. 3.4e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```

```

OY 1 VLXDDLE 8
DB 201 TLDDLE 208

```

```

RESULT 29
ABB21583
ID ABB21583 standard; Protein: 222 AA.
AC ABB21583;
XX 18-FEB-2002 (first entry)
DT
XX
XX Novel human diagnostic protein #21574.
DE
XX
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
XX Homo sapiens.
OS
XX
XX WO200175067-A2.
PN
XX
XX 11-OCT-2001.
PD
XX
XX 30-MAR-2001; 2001WO-US08631.
PF
XX
XX 31-MAR-2000; 2000US-0540217.
PR
XX
XX 23-AUG-2000; 2000US-0649167.
XX
XX (HYSE-) HYSEQ INC.
PA
XX
XX Drmanac RT, Liu C, Tang YT;
PI
XX
XX WPI; 2001-639362/73.
DR
XX
XX N-PSDB; AAS85770.
XX
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity
XX
XX Claim 20; SEQ ID NO 51942; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC

```

CC for identifying expressed genes. (I) is useful in gene therapy techniques  
CC to restore normal activity of (II) or to treat disease states involving  
CC (II). (II) is useful for generating antibodies against it, detecting or  
CC quantitating a polypeptide in tissue, as molecular weight markers and as  
CC a food supplement. (II) and its binding partners are useful in medical  
CC imaging of sites expressing (II). (I) and (II) are useful for treating  
CC disorders involving aberrant protein expression or biological activity.  
CC The polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. ABG00010-ABG30377 represent novel human  
CC diagnostic amino acid sequences of the invention.  
CC Note: The sequence data for this patent did not appear in the printed  
CC specification, but was obtained in electronic format directly from WIP0  
CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 222 AA;

Query Match 76.3%; Score 29; DB 22; Length 222;  
Best Local Similarity 55.6%; Pred. No. 3.5e+02;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLXDDLEA 9  
| : ||: ||  
Db 143 VINDVIEA 151

RESULT 30  
AAG57749  
ID AAG57749 standard; Protein: 224 AA.  
XX  
XX AAG57749;

DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SMO ID NO: 74457.

KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
XX termination sequence.

OS Arabidopsis thaliana.

PN EP1033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 04-MAY-1999; 99US-0132407.  
PR 05-MAY-1999; 99US-0132484.  
PR 06-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132487.  
PR 11-MAY-1999; 99US-0132863.  
PR 99US-0134256.

PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135533.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139492.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139896.  
PR 23-JUN-1999; 99US-0140353.  
PR 24-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 21-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 22-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.

```

PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145951.
PR 28-JUL-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 06-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147933.
PR 09-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 04-OCT-1999; 99US-0157733.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159584.
PR 18-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.

```

```

PR 22-OCT-1999; 99US-0160989.
PR 23-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161921.
PR 28-OCT-1999; 99US-0161993.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

```

Query Match 76.3%; Score 29; DB 21; Length 224;
Best Local Similarity 75.0%; Pred. No. 3.5e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 VLXDDLE 8
DB 9 LKIDLE 16

```

## RESULT 31

```

ID ABB54117 standard; Protein; 244 AA.
XX ABB54117;

```

```

AC ABB54117;

```

```

DT 16-MAY-2002 (first entry)

```

```

DE Lactococcus lactis protein ylbP.

```

```

XX Biosynthesis; biodegradation; lactic bacterium; yogurt; cheese.

```

```

OS Lactococcus lactis IL1403.

```

```

XX FR2807446-A1.

```

```

PD 12-OCT-2001.

```

```

PF 11-APR-2000; 2000FR-0004630.

```

```

XX 11-APR-2000; 2000FR-0004630.

```

```

PA (INRG) INRA INST NAT RECH AGRONOMIQUE.

```

```

PI Bolotine A, Sorokine A, Renault P, Ehrlich SD;

```

```

DR WPI; 2002-043418/06.

```

```

XX New nucleotide sequence useful in the identification of Lactococcus

```

```

PT lactis and related species -

```

```

PS Claim 6; SEQ ID No 819; 2504pp; French.

```

```

XX The present invention is related to a Lactococcus lactis nucleotide
CC sequence (AB90521) and related proteins (AB53300-AB55621). The
CC nucleic acid sequence is useful in the detection and/or amplification of
CC nucleic acid sequence, particularly to identify Lactococcus lactis or
CC related species. The proteins of the invention are useful for the
CC biosynthesis or biodegradation of a composition of interest. The
CC invention helps research in lactic bacteria, particularly useful in the
CC production of yogurt and cheese.

```

```

CC Note: The sequence data for this patent is based on equivalent patent
CC WO20017734 (published 18-OCT-2001) which is available in electronic
CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

```

```

XX Sequence 244 AA;

```

```

SQ Query Match 76.3%; Score 29; DB 23; Length 244;

```

```

Best Local Similarity 100.0%; Pred. No. 3.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 4 DDLEA 9

```



Db 71 DDLLEA 76  
|||||  
RESULT 32  
AAG57748  
ID AAG57748 standard; Protein: 249 AA.  
XX AAG57748;  
AC  
XX 18-OCT-2000 (first entry)  
DT  
XX Arabidopsis thaliana protein fragment SEQ ID NO: 74456.  
DE  
XX Protein identification: signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
XX Arabidopsis thaliana.  
XX  
XX EPI033405-A2.  
XX  
PD 06-SEP-2000.  
XX  
XX 25-FEB-2000; 2000EP-0301439.  
XX  
XX 25-FEB-1999; 990S-0121825.  
PR 05-MAR-1999; 990S-0123180.  
PR 09-MAR-1999; 990S-0123548.  
PR 23-MAR-1999; 990S-0125788.  
PR 25-MAR-1999; 990S-0126264.  
PR 29-MAR-1999; 990S-0126785.  
PR 01-APR-1999; 990S-0127462.  
PR 06-APR-1999; 990S-0128234.  
PR 08-APR-1999; 990S-0128714.  
PR 16-APR-1999; 990S-0129845.  
PR 19-APR-1999; 990S-0130077.  
PR 21-APR-1999; 990S-0130449.  
PR 23-APR-1999; 990S-0130510.  
PR 23-APR-1999; 990S-0130891.  
PR 28-APR-1999; 990S-0131449.  
PR 30-APR-1999; 990S-0132048.  
PR 30-APR-1999; 990S-0132407.  
PR 04-MAY-1999; 990S-0132484.  
PR 05-MAY-1999; 990S-0132485.  
PR 06-MAY-1999; 990S-0132486.  
PR 06-MAY-1999; 990S-0132487.  
PR 07-MAY-1999; 990S-0132863.  
PR 11-MAY-1999; 990S-0134256.  
PR 14-MAY-1999; 990S-0134218.  
PR 14-MAY-1999; 990S-0134219.  
PR 14-MAY-1999; 990S-0134221.  
PR 14-MAY-1999; 990S-0134370.  
PR 18-MAY-1999; 990S-0134768.  
PR 19-MAY-1999; 990S-0134941.  
PR 20-MAY-1999; 990S-0135124.  
PR 21-MAY-1999; 990S-0135353.  
PR 24-MAY-1999; 990S-0135629.  
PR 25-MAY-1999; 990S-0136021.  
PR 27-MAY-1999; 990S-0136382.  
PR 28-MAY-1999; 990S-0136782.  
PR 01-JUN-1999; 990S-0137222.  
PR 03-JUN-1999; 990S-0137528.  
PR 04-JUN-1999; 990S-0137502.  
PR 07-JUN-1999; 990S-0137724.  
PR 08-JUN-1999; 990S-0138094.  
PR 10-JUN-1999; 990S-0138540.  
PR 14-JUN-1999; 990S-0138847.  
PR 16-JUN-1999; 990S-0139119.  
PR 16-JUN-1999; 990S-0139452.  
PR 17-JUN-1999; 990S-0139453.  
PR 18-JUN-1999; 990S-0139492.  
PR 18-JUN-1999; 990S-0139454.

PR 18-JUN-1999; 990S-0139455.  
PR 18-JUN-1999; 990S-0139456.  
PR 18-JUN-1999; 990S-0139457.  
PR 18-JUN-1999; 990S-0139458.  
PR 18-JUN-1999; 990S-0139459.  
PR 18-JUN-1999; 990S-0139460.  
PR 18-JUN-1999; 990S-0139461.  
PR 18-JUN-1999; 990S-0139462.  
PR 18-JUN-1999; 990S-0139463.  
PR 18-JUN-1999; 990S-0139750.  
PR 18-JUN-1999; 990S-0139751.  
PR 21-JUN-1999; 990S-0139817.  
PR 22-JUN-1999; 990S-0139899.  
PR 23-JUN-1999; 990S-0140353.  
PR 23-JUN-1999; 990S-0140354.  
PR 24-JUN-1999; 990S-0140695.  
PR 28-JUN-1999; 990S-0140823.  
PR 28-JUN-1999; 990S-0140991.  
PR 30-JUN-1999; 990S-0141287.  
PR 01-JUL-1999; 990S-0141842.  
PR 02-JUL-1999; 990S-0142154.  
PR 06-JUL-1999; 990S-0142055.  
PR 08-JUL-1999; 990S-0142390.  
PR 09-JUL-1999; 990S-0142803.  
PR 12-JUL-1999; 990S-0142920.  
PR 13-JUL-1999; 990S-0142977.  
PR 14-JUL-1999; 990S-0143542.  
PR 15-JUL-1999; 990S-0143624.  
PR 16-JUL-1999; 990S-0144005.  
PR 16-JUL-1999; 990S-0144086.  
PR 19-JUL-1999; 990S-0144086.  
PR 19-JUL-1999; 990S-0144325.  
PR 19-JUL-1999; 990S-0144331.  
PR 19-JUL-1999; 990S-0144332.  
PR 19-JUL-1999; 990S-0144333.  
PR 19-JUL-1999; 990S-0144334.  
PR 20-JUL-1999; 990S-0144335.  
PR 20-JUL-1999; 990S-0144352.  
PR 20-JUL-1999; 990S-0144632.  
PR 21-JUL-1999; 990S-0144884.  
PR 21-JUL-1999; 990S-0144814.  
PR 21-JUL-1999; 990S-0145086.  
PR 22-JUL-1999; 990S-0145088.  
PR 22-JUL-1999; 990S-0145085.  
PR 22-JUL-1999; 990S-0145087.  
PR 22-JUL-1999; 990S-0145089.  
PR 23-JUL-1999; 990S-0145192.  
PR 23-JUL-1999; 990S-0145145.  
PR 23-JUL-1999; 990S-0145218.  
PR 26-JUL-1999; 990S-0145224.  
PR 27-JUL-1999; 990S-0145276.  
PR 27-JUL-1999; 990S-0145913.  
PR 27-JUL-1999; 990S-0145918.  
PR 28-JUL-1999; 990S-0145919.  
PR 02-AUG-1999; 990S-0145951.  
PR 02-AUG-1999; 990S-0146386.  
PR 02-AUG-1999; 990S-0146388.  
PR 03-AUG-1999; 990S-0146389.  
PR 04-AUG-1999; 990S-0147038.  
PR 04-AUG-1999; 990S-0147204.  
PR 05-AUG-1999; 990S-0147302.  
PR 05-AUG-1999; 990S-0147192.  
PR 06-AUG-1999; 990S-0147260.  
PR 06-AUG-1999; 990S-0147303.  
PR 09-AUG-1999; 990S-0147416.  
PR 09-AUG-1999; 990S-0147493.  
PR 10-AUG-1999; 990S-0147935.  
PR 11-AUG-1999; 990S-0148171.  
PR 12-AUG-1999; 990S-0148319.  
PR 13-AUG-1999; 990S-0148341.  
PR 13-AUG-1999; 990S-0148565.  
PR 16-AUG-1999; 990S-0148684.  
PR 16-AUG-1999; 990S-0149368.

```

PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

```

Query Match          76.3%; Score 29; DB 21; Length 249;
Best Local Similarity 75.0%; Pred. No. 4e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy 1 VLXDDLE 8
   : |||||
Db 34 LLDLLE 41

```

```

RESULT 33
AAB60860
ID AAB60860 standard; protein: 292 AA.

```

```

XX AAB60860;
XX 30-MAR-2001 (first entry)
XX Mevalonate pathway protein #3.
XX Mevalonate pathway; disease; infection.
XX Streptococcus pyogenes.
XX WO200078935-A1.
XX 28-DEC-2000.
XX 22-JUN-2000; 2000WO-US17262.
XX 22-JUN-1999; 99US-0140519.
XX 02-AUG-1999; 99US-0146682.
XX (SMIR ) SMITHKLINE BEECHAM CORP.
XX (SMIR ) SMITHKLINE BEECHAM PLC.
XX Brown JR, Gwynn M, Machie TB, Myers JE, Traini CM, Van Horn S;
XX Wilding EI;
XX WPI; 2001-071392/08.
XX PT New isolated mevalonate pathway gene polynucleotide derived from
XX bacterium is useful for treatment of bacterial infection -
XX Claim 20; Page 14; 158pp; English.
XX CC The present invention relates to an isolated mevalonate pathway gene
XX derived from a bacterium from clade of Class II of the
XX phylogenetic tree referred to in the specification. The invention may be
XX used for treatment of disease related to bacterial infection, e.g.
XX conjunctivitis, pneumonia, bacteraemia and meningitis.
XX SQ Sequence 292 AA;

```

```

Query Match          76.3%; Score 29; DB 22; Length 292;
Best Local Similarity 85.7%; Pred. No. 4.7e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy 2 LXDLE 8
   : |||||
Db 111 LLDLLE 117

```

```

RESULT 34
ABP28408
ID ABP28408 standard; protein: 297 AA.
XX
XX ABP28408;

```

```

XX 02-JUL-2002 (first entry)
XX Streptococcus polypeptide SEQ ID NO 5992.
XX
XX Streptococcus polypeptide SEQ ID NO 5992.
XX
XX Streptococcus GAS; GAS; group B streptococcus; Streptococcus agalactiae;
XX group A streptococcus; Streptococcus pyogenes; antibacterial;
XX antiinflammatory; infection; vaccine; meningitis; gene therapy.
XX Streptococcus pyogenes.
XX
XX WO200234771-A2.
XX
XX 02-MAY-2002.
XX
XX 29-OCT-2001; 2001WO-GB04789.
XX
XX 27-OCT-2000; 2000GB-0026333.

```

PR 24-NOV-2000; 2000GB-0028727.  
PR 07-MAR-2001; 2001GB-0005640.  
XX  
XX  
PA (CHIR-) CHIRON SPA.  
PA (GENO-) INST GENOMIC RES.  
XX  
XX  
PI Telford J, Masiagnani V, Margalit Ros YI, Grandi G, Fraser C;  
PI Tectelin H;  
XX  
XX  
DR WPI: 2002-352536/38.  
DR N-PSDB: ABN69039.  
XX  
XX  
PR New Streptococcus protein for the treatment or prevention of infection  
PT or disease caused by Streptococcus bacteria, such as meningitis, and  
PS for detecting a compound that binds to the protein -  
PS Claim 1; Page 3761; 4525pp; English.  
XX  
XX  
CC The invention relates to a protein (ABP25413-ABP30895) from group B  
CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS  
CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in  
CC the specification. The proteins have antibacterial and antiinflammatory  
CC activity. (I), nucleic acids encoding (I), ABN60044-ABN71526 and  
CC antibodies that bind (I) are used in the manufacture of medicaments for  
CC the treatment or prevention of infection or disease caused by  
CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.  
CC Nucleic acids encoding (I) are used to detect Streptococcus in a  
CC biological sample. (I) is used to determine whether a compound binds to  
CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be  
CC used as a vaccine or diagnostic composition. The disease caused by  
CC Streptococcus that is prevented or treated may be meningitis. Nucleic  
CC acid encoding (I) may be used to recombinantly produce (I) and may be  
CC used in gene therapy. Antibodies to (I) are used for affinity  
CC chromatography, immunoassays, and distinguishing/identifying  
CC Streptococcus proteins.  
XX  
XX  
SQ Sequence 297 AA:  
  
Query Match 76.3%; Score 29; DB 23; Length 297;  
Best Local Similarity 85.7%; Pred. No. 4.8e+02;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 2 LXDLL 8  
I I I I I  
Db 116 LSDLL 122  
  
RESULT 35  
AAG04826  
ID AAG04826 standard; Protein; 358 AA.  
XX  
XX  
AC AAG04826;  
XX  
XX  
DI 17-OCT-2000 (first entry)  
XX  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 999.  
XX  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hydridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX  
XX  
OS Arabidopsis thaliana.  
XX  
XX  
PN EP1033405-A2.  
PD  
XX  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
XX  
PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 28-APR-1999; 99US-0130891.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132487.  
PR 11-MAY-1999; 99US-0132863.  
PR 14-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 18-MAY-1999; 99US-0134370.  
PR 19-MAY-1999; 99US-0134768.  
PR 20-MAY-1999; 99US-0134941.  
PR 21-MAY-1999; 99US-0135124.  
PR 24-MAY-1999; 99US-0135353.  
PR 25-MAY-1999; 99US-0135629.  
PR 27-MAY-1999; 99US-0136021.  
PR 28-MAY-1999; 99US-0136392.  
PR 01-JUN-1999; 99US-0136782.  
PR 03-JUN-1999; 99US-0137222.  
PR 04-JUN-1999; 99US-0137528.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 14-JUN-1999; 99US-0138847.  
PR 16-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 17-JUN-1999; 99US-0139453.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139819.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140685.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143442.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.



PR	21-APR-1999;	99US-0130449.
PR	23-APR-1999;	99US-0130510.
PR	23-APR-1999;	99US-0130891.
PR	28-APR-1999;	99US-0131449.
PR	30-APR-1999;	99US-0132048.
PR	04-MAY-1999;	99US-0132407.
PR	05-MAY-1999;	99US-0132464.
PR	06-MAY-1999;	99US-0132486.
PR	07-MAY-1999;	99US-0132487.
PR	11-MAY-1999;	99US-0132863.
PR	14-MAY-1999;	99US-0134218.
PR	14-MAY-1999;	99US-0134221.
PR	14-MAY-1999;	99US-0134370.
PR	18-MAY-1999;	99US-0134768.
PR	19-MAY-1999;	99US-0134941.
PR	20-MAY-1999;	99US-0135124.
PR	21-MAY-1999;	99US-0135353.
PR	24-MAY-1999;	99US-0136021.
PR	27-MAY-1999;	99US-0136392.
PR	28-MAY-1999;	99US-0136782.
PR	01-JUN-1999;	99US-0137222.
PR	03-JUN-1999;	99US-0137528.
PR	04-JUN-1999;	99US-0137502.
PR	07-JUN-1999;	99US-0137724.
PR	08-JUN-1999;	99US-0138094.
PR	10-JUN-1999;	99US-0138540.
PR	10-JUN-1999;	99US-0138847.
PR	14-JUN-1999;	99US-0139119.
PR	16-JUN-1999;	99US-0139453.
PR	17-JUN-1999;	99US-0139492.
PR	18-JUN-1999;	99US-0139454.
PR	18-JUN-1999;	99US-0139455.
PR	18-JUN-1999;	99US-0139456.
PR	18-JUN-1999;	99US-0139457.
PR	18-JUN-1999;	99US-0139458.
PR	18-JUN-1999;	99US-0139459.
PR	18-JUN-1999;	99US-0139460.
PR	18-JUN-1999;	99US-0139461.
PR	18-JUN-1999;	99US-0139462.
PR	18-JUN-1999;	99US-0139463.
PR	18-JUN-1999;	99US-0139750.
PR	21-JUN-1999;	99US-0139763.
PR	22-JUN-1999;	99US-0139817.
PR	23-JUN-1999;	99US-0139899.
PR	23-JUN-1999;	99US-0140353.
PR	24-JUN-1999;	99US-0140695.
PR	28-JUN-1999;	99US-0140823.
PR	29-JUN-1999;	99US-0140991.
PR	30-JUN-1999;	99US-0141287.
PR	01-JUL-1999;	99US-0141842.
PR	01-JUL-1999;	99US-0142154.
PR	02-JUL-1999;	99US-0142055.
PR	06-JUL-1999;	99US-0142390.
PR	08-JUL-1999;	99US-0142803.
PR	09-JUL-1999;	99US-0142920.
PR	12-JUL-1999;	99US-0142977.
PR	13-JUL-1999;	99US-0143542.
PR	14-JUL-1999;	99US-0143624.
PR	15-JUL-1999;	99US-0144005.
PR	16-JUL-1999;	99US-0144085.
PR	16-JUL-1999;	99US-0144086.
PR	19-JUL-1999;	99US-0144325.
PR	19-JUL-1999;	99US-0144331.
PR	19-JUL-1999;	99US-0144332.
PR	19-JUL-1999;	99US-0144333.
PR	19-JUL-1999;	99US-0144334.
PR	19-JUL-1999;	99US-0144335.
<hr/>		
PR	20-JUL-1999;	99US-0144352.
PR	20-JUL-1999;	99US-0144632.
PR	21-JUL-1999;	99US-0144884.
PR	21-JUL-1999;	99US-0144814.
PR	21-JUL-1999;	99US-0145086.
PR	22-JUL-1999;	99US-0145087.
PR	22-JUL-1999;	99US-0145087.
PR	22-JUL-1999;	99US-0145089.
PR	23-JUL-1999;	99US-0145145.
PR	23-JUL-1999;	99US-0145224.
PR	26-JUL-1999;	99US-0145276.
PR	27-JUL-1999;	99US-0145913.
PR	27-JUL-1999;	99US-0145918.
PR	27-JUL-1999;	99US-0145919.
PR	28-JUL-1999;	99US-0145951.
PR	02-AUG-1999;	99US-0146386.
PR	02-AUG-1999;	99US-0146388.
PR	03-AUG-1999;	99US-0147038.
PR	04-AUG-1999;	99US-0147204.
PR	05-AUG-1999;	99US-0147302.
PR	05-AUG-1999;	99US-0147192.
PR	06-AUG-1999;	99US-0147260.
PR	06-AUG-1999;	99US-0147303.
PR	09-AUG-1999;	99US-0147416.
PR	09-AUG-1999;	99US-0147493.
PR	10-AUG-1999;	99US-0147935.
PR	11-AUG-1999;	99US-0148171.
PR	12-AUG-1999;	99US-0148319.
PR	13-AUG-1999;	99US-0148565.
PR	13-AUG-1999;	99US-0148684.
PR	16-AUG-1999;	99US-0149368.
PR	17-AUG-1999;	99US-0149175.
PR	18-AUG-1999;	99US-0149426.
PR	20-AUG-1999;	99US-0149722.
PR	20-AUG-1999;	99US-0149723.
PR	20-AUG-1999;	99US-0149724.
PR	23-AUG-1999;	99US-0149929.
PR	23-AUG-1999;	99US-0149930.
PR	25-AUG-1999;	99US-0150566.
PR	26-AUG-1999;	99US-0150884.
PR	27-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-0151066.
PR	27-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151303.
PR	31-AUG-1999;	99US-0151438.
PR	01-SEP-1999;	99US-0151930.
PR	07-SEP-1999;	99US-0152363.
PR	10-SEP-1999;	99US-0153070.
PR	13-SEP-1999;	99US-0153758.
PR	15-SEP-1999;	99US-0154018.
PR	16-SEP-1999;	99US-0154039.
PR	20-SEP-1999;	99US-0154779.
PR	22-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-0155486.
PR	24-SEP-1999;	99US-0155659.
PR	28-SEP-1999;	99US-0156458.
PR	29-SEP-1999;	99US-0156596.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.

PR	14-OCT-1999	99US-01593331
PR	14-OCT-1999	99US-01596337
PR	14-OCT-1999	99US-01596384
PR	18-OCT-1999	99US-01595838
PR	21-OCT-1999	99US-01607411
PR	21-OCT-1999	99US-01607647
PR	21-OCT-1999	99US-01607658
PR	21-OCT-1999	99US-01607677
PR	21-OCT-1999	99US-01608140
PR	21-OCT-1999	99US-01608174
PR	21-OCT-1999	99US-01608175
PR	21-OCT-1999	99US-01608185
PR	22-OCT-1999	99US-01609880
PR	22-OCT-1999	99US-01609881
PR	22-OCT-1999	99US-01609882
PR	25-OCT-1999	99US-01614004
PR	25-OCT-1999	99US-01614005
PR	25-OCT-1999	99US-01614006
PR	26-OCT-1999	99US-01613550
PR	26-OCT-1999	99US-01613569
PR	26-OCT-1999	99US-01613621
PR	28-OCT-1999	99US-01619320
PR	28-OCT-1999	99US-01619322
PR	28-OCT-1999	99US-01619393
PR	29-OCT-1999	99US-01621422

Query Match	76.3%	Score 29;	DB 21;	Length 358;
Best Local Similarity	75.0%	Pred No.	5.8e+02;	
Matches	6;	Conservative	1;	Indels 0; Caps 0

QY	1 VLXDDLE	8
	:	
Db	9 LLIDDLE	16

RESULT 37  
ABG29716  
ID ABG29716 standard; Protein; 396 AA

AC ABG29716;

DT 18-FEB-2002 (first entry)

DE Novel human diagnostic protein #29707.

**KW** Human; chromosome mapping; gene mapping; gene therapy; forensic food supplement; medical imaging; diagnostic; genetic disorder.

05 Homo sapiens

PN WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217

XX  
XX  
R (UNCLAS) NUCO TNC

XX  
4  
3

PI Drmanac RT, Liu C, Yang Y, et al. 2000. A

DR WPL; 2001-639364/13.  
DR N-PSDB: AAS93903.

XX  
XX  
PT

New isolated poly

PT New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity -

PS Claim 20; SEQ ID No 60075; 103pp; English

CC The invention relates to isolated polynucleotide (I) and  
CC polypeptide (II) sequences. (I) is useful as hybridisation probes

polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. ABC00010-ABC3037 represent novel human diagnostic amino acid sequences of the invention.

Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at [http://wipo.int/pub/published\\_pct](http://wipo.int/pub/published_pct) sequences.

Sequence 396 AA

Query Match:	Score 29; DB 22; Length 396;
Best Local Similarity	75.0%; Pred No. 6.4e+02;
Matches	6; Conservative 1; Indels 0; Gaps 0;

```
QY      2  LXDDLEA  9
          |  | | | |
Db      252  LGDDVLEA 259
```

RESULT 38  
AAB94128  
ID AAB94128 standard; Protein; 405 AA

AC AAB94128;

DT 26-JUN-2001 (first entry)

DE Human protein sequence SEQ ID NO:14383.

**KW** Human; primer; detection; diagnosis; antisense therapy; gene therapy.

OS Homo sapiens.

PN EP1074617-A2.

PD 07-FEB-2001

28-JUL-2000; 2000EP-0116126-

PR	29-JUL-1999;	99JP-0248036.
PR	27-AUG-1999.	99JP-0300253

PR 11-JAN-2000; 2000JP-0118776.  
PR 03-MAY-2000; 2000JP-0183767

PR 09-JUN-2000; 2000JP-0241899.  
XX XX

PA (HELI-) HELIX RES INST.  
YY

PI	Ota T,	Isogai T,	Nishikawa
PI	Ishii S	Suoiyama T	Wakam

XX DB WPT: 2001-318749/34.

XX	primer sets for synt
PT	

PT and/or diagnosis of the abnormality of the proteins encoded by the full-length CDNAs derived in the specification, and for the selection

PT. TULL-JENGLI CDNAS -  
XX

AA	PS	Claim ID	SEQ ID	pp	CD ROM	English
		8;	14383;	2537		

The present invention describes primer sets for synthesising 5002  
CC full-length cDNAs defined in the specification. Where a primer set  
CC

CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary  
CC to the complementary strand of a polynucleotide which comprises one of  
CC the 5602 nucleotide sequences defined in the specification, where the  
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
CC of an oligonucleotide comprising a sequence complementary to the  
CC sequence and an oligonucleotide comprising a sequence complementary to a  
CC polynucleotide which comprises a 3'-end sequence, where the  
CC oligonucleotide comprises at least 15 nucleotides and the combination of  
CC the 5'-end sequence/3'-end sequence is selected from those defined in  
CC the specification. The primer sets can be used in antisense therapy and  
CC in gene therapy. The primers are useful for synthesising polynucleotides,  
CC particularly full-length cDNAs. The primers are also useful for the  
CC detection and/or diagnosis of the abnormality of the full-length  
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to  
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632  
CC represent oligonucleotides, all of which are used in the exemplification  
CC of the present invention.

XX Sequence 405 AA;

Query Match 76.38; Score 29; DB 22; Length 405;

Best Local Similarity 66.78; Pred. No. 6.6e+02;

Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 VLXDDLEA 9

Db 153 ILDDLSEA 161

RESULT 39

AAG04825 AAG04825 standard; Protein: 414 AA.

XX AAG04825;

XX 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 998.

XX Protein identification; signal transduction pathway; metabolic pathway;  
KM hybridisation assay; genetic mapping; gene expression control; promoter;  
KM termination sequence.

XX Arabidopsis thaliana.

OS EPI033405-A2.

PN 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0122548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0128845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130410.

PR 23-APR-1999; 99US-0130510.

PR 28-APR-1999; 99US-0130891.

PR 30-APR-1999; 99US-0132048.

PR 04-MAY-1999; 99US-0132407.

PR 05-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134914.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 17-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139753.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142380.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144684.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.



```

PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 30-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 23-SEP-1999; 99US-0155659.
PR 24-SEP-1999; 99US-0156458.
PR 28-SEP-1999; 99US-0156596.
PR 29-SEP-1999; 99US-0157117.
PR 04-OCT-1999; 99US-0157753.
PR 05-OCT-1999; 99US-0157865.
PR 06-OCT-1999; 99US-0158023.
PR 07-OCT-1999; 99US-0158023.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 12-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 13-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 16-OCT-1999; 99US-0159584.
PR 18-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.

```

```

PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161044.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162142.

Query Match 76.3%; Score 29; DB 21; Length 414;
Best Local Similarity 75.0%; Pred. No. 6.7e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VKDDLE 8
Db 65 LKDDLE 72

RESULT 40
AAC61478
ID AAC61478 standard; protein; 414 AA.
XX AAC61478;
AC AAC61478;
XX 18-OCT-2000 (first entry)
DT
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 79745.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
OS Arabidopsis thaliana.
XX EP1033405-A2.
PN
XX
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123160.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0128845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134221.

```

PR 14-MAY-1999; 990S-0134370.  
PR 18-MAY-1999; 990S-0134768.  
PR 19-MAY-1999; 990S-0134941.  
PR 20-MAY-1999; 990S-0135124.  
PR 21-MAY-1999; 990S-0135523.  
PR 24-MAY-1999; 990S-0135629.  
PR 25-MAY-1999; 990S-0136021.  
PR 27-MAY-1999; 990S-0136392.  
PR 28-MAY-1999; 990S-0136782.  
PR 01-JUN-1999; 990S-0137222.  
PR 03-JUN-1999; 990S-0137528.  
PR 04-JUN-1999; 990S-0137502.  
PR 07-JUN-1999; 990S-0137724.  
PR 08-JUN-1999; 990S-0138094.  
PR 10-JUN-1999; 990S-0138540.  
PR 14-JUN-1999; 990S-0138847.  
PR 16-JUN-1999; 990S-0139119.  
PR 16-JUN-1999; 990S-0139452.  
PR 17-JUN-1999; 990S-0139453.  
PR 18-JUN-1999; 990S-0139492.  
PR 18-JUN-1999; 990S-0139454.  
PR 18-JUN-1999; 990S-0139455.  
PR 18-JUN-1999; 990S-0139456.  
PR 18-JUN-1999; 990S-0139457.  
PR 18-JUN-1999; 990S-0139458.  
PR 18-JUN-1999; 990S-0139459.  
PR 18-JUN-1999; 990S-0139460.  
PR 18-JUN-1999; 990S-0139461.  
PR 18-JUN-1999; 990S-0139462.  
PR 18-JUN-1999; 990S-0139463.  
PR 18-JUN-1999; 990S-0139750.  
PR 21-JUN-1999; 990S-0139763.  
PR 22-JUN-1999; 990S-0139817.  
PR 23-JUN-1999; 990S-0139899.  
PR 23-JUN-1999; 990S-0140353.  
PR 24-JUN-1999; 990S-0140354.  
PR 28-JUN-1999; 990S-0140695.  
PR 29-JUN-1999; 990S-0140823.  
PR 30-JUN-1999; 990S-0140991.  
PR 01-JUL-1999; 990S-0141287.  
PR 01-JUL-1999; 990S-0141842.  
PR 02-JUL-1999; 990S-0142154.  
PR 06-JUL-1999; 990S-0142055.  
PR 08-JUL-1999; 990S-0142390.  
PR 09-JUL-1999; 990S-0142803.  
PR 12-JUL-1999; 990S-0142920.  
PR 13-JUL-1999; 990S-0143542.  
PR 14-JUL-1999; 990S-0143624.  
PR 15-JUL-1999; 990S-0144005.  
PR 16-JUL-1999; 990S-0144085.  
PR 16-JUL-1999; 990S-0144086.  
PR 19-JUL-1999; 990S-0144325.  
PR 19-JUL-1999; 990S-0144331.  
PR 19-JUL-1999; 990S-0144332.  
PR 19-JUL-1999; 990S-0144333.  
PR 19-JUL-1999; 990S-0144334.  
PR 20-JUL-1999; 990S-0144335.  
PR 20-JUL-1999; 990S-0144352.  
PR 20-JUL-1999; 990S-0144632.  
PR 21-JUL-1999; 990S-0144884.  
PR 21-JUL-1999; 990S-0144814.  
PR 21-JUL-1999; 990S-0145086.  
PR 22-JUL-1999; 990S-0145088.  
PR 22-JUL-1999; 990S-0145085.  
PR 22-JUL-1999; 990S-0145087.  
PR 22-JUL-1999; 990S-0145089.  
PR 23-JUL-1999; 990S-0145192.  
PR 23-JUL-1999; 990S-0145145.  
PR 23-JUL-1999; 990S-0145218.  
PR 26-JUL-1999; 990S-0145224.  
PR 27-JUL-1999; 990S-0145276.  
PR 27-JUL-1999; 990S-0145913.

---

PR 27-JUL-1999; 990S-0145918.  
PR 27-JUL-1999; 990S-0145919.  
PR 28-JUL-1999; 990S-0145951.  
PR 02-AUG-1999; 990S-0146386.  
PR 02-AUG-1999; 990S-0146388.  
PR 03-AUG-1999; 990S-0146389.  
PR 04-AUG-1999; 990S-0147038.  
PR 04-AUG-1999; 990S-0147204.  
PR 05-AUG-1999; 990S-0147302.  
PR 05-AUG-1999; 990S-0147192.  
PR 06-AUG-1999; 990S-0147260.  
PR 06-AUG-1999; 990S-0147303.  
PR 09-AUG-1999; 990S-0147416.  
PR 09-AUG-1999; 990S-0147935.  
PR 10-AUG-1999; 990S-0148171.  
PR 11-AUG-1999; 990S-0148319.  
PR 12-AUG-1999; 990S-0148341.  
PR 13-AUG-1999; 990S-0148565.  
PR 13-AUG-1999; 990S-0148684.  
PR 16-AUG-1999; 990S-0149368.  
PR 17-AUG-1999; 990S-0149370.  
PR 18-AUG-1999; 990S-0149175.  
PR 20-AUG-1999; 990S-0149426.  
PR 20-AUG-1999; 990S-0149722.  
PR 20-AUG-1999; 990S-0149723.  
PR 20-AUG-1999; 990S-0149929.  
PR 23-AUG-1999; 990S-0149902.  
PR 23-AUG-1999; 990S-0149930.  
PR 25-AUG-1999; 990S-0150566.  
PR 26-AUG-1999; 990S-0150884.  
PR 27-AUG-1999; 990S-0151065.  
PR 27-AUG-1999; 990S-0151066.  
PR 27-AUG-1999; 990S-0151080.  
PR 30-AUG-1999; 990S-0151303.  
PR 31-AUG-1999; 990S-0151438.  
PR 01-SEP-1999; 990S-0151930.  
PR 07-SEP-1999; 990S-0152363.  
PR 10-SEP-1999; 990S-0153070.  
PR 13-SEP-1999; 990S-0153758.  
PR 15-SEP-1999; 990S-0154018.  
PR 16-SEP-1999; 990S-0154039.  
PR 20-SEP-1999; 990S-0154779.  
PR 22-SEP-1999; 990S-0155139.  
PR 23-SEP-1999; 990S-0155486.  
PR 24-SEP-1999; 990S-0155659.  
PR 28-SEP-1999; 990S-0156458.  
PR 29-SEP-1999; 990S-0156596.  
PR 04-OCT-1999; 990S-0157117.  
PR 05-OCT-1999; 990S-0157753.  
PR 06-OCT-1999; 990S-0157865.  
PR 07-OCT-1999; 990S-0158029.  
PR 08-OCT-1999; 990S-0158232.  
PR 12-OCT-1999; 990S-0158369.  
PR 13-OCT-1999; 990S-0159293.  
PR 13-OCT-1999; 990S-0159294.  
PR 13-OCT-1999; 990S-0159295.  
PR 14-OCT-1999; 990S-0159329.  
PR 14-OCT-1999; 990S-0159330.  
PR 14-OCT-1999; 990S-0159331.  
PR 14-OCT-1999; 990S-0159637.  
PR 14-OCT-1999; 990S-0159638.  
PR 18-OCT-1999; 990S-0159584.  
PR 21-OCT-1999; 990S-0160741.  
PR 21-OCT-1999; 990S-0160767.  
PR 21-OCT-1999; 990S-0160768.  
PR 21-OCT-1999; 990S-0160770.  
PR 21-OCT-1999; 990S-0160814.  
PR 22-OCT-1999; 990S-0160815.  
PR 22-OCT-1999; 990S-0160980.  
PR 22-OCT-1999; 990S-0160981.  
PR 22-OCT-1999; 990S-0160989.  
PR 25-OCT-1999; 990S-0161404.  
PR 25-OCT-1999; 990S-0161405.

PR 25-OCT-1999; 99US-0161406.  
PR 26-OCT-1999; 99US-0161359.  
PR 26-OCT-1999; 99US-0161360.  
PR 26-OCT-1999; 99US-0161361.  
PR 28-OCT-1999; 99US-0161920.  
PR 28-OCT-1999; 99US-0161921.  
PR 28-OCT-1999; 99US-0161993.  
PR 29-OCT-1999; 99US-0162142.  
Query Match 76.3%; Score 29; DB 21; Length 414;  
Best Local Similarity 75.0%; Pred. No. 6.7e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 1 VLXDDLE 8  
DB 65 LIDDDLE 72  
RESULT 41  
AAG04824  
ID AAG04824 standard; Protein: 415 AA.  
AC AAG04824;  
XX  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 997.  
XX  
XX Protein identification; signal transduction pathway; metabolic pathway;  
XX KM hybridisation assay; genetic mapping; gene expression control; promoter;  
XX KW termination sequence.  
XX OS Arabidopsis thaliana.  
XX PN EP1033405-A2.  
XX  
XX 06-SEP-2000.  
XX  
XX 25-FEB-2000; 2000EP-0301439.  
XX  
XX 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123180.  
PR 09-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0125788.  
PR 25-MAR-1999; 99US-0126284.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 19-APR-1999; 99US-0130077.  
PR 21-APR-1999; 99US-0130449.  
PR 23-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0132048.  
PR 30-APR-1999; 99US-0132407.  
PR 04-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132485.  
PR 06-MAY-1999; 99US-0132486.  
PR 07-MAY-1999; 99US-0132487.  
PR 07-MAY-1999; 99US-0132863.  
PR 11-MAY-1999; 99US-0134256.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134221.  
PR 14-MAY-1999; 99US-0134370.  
PR 18-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 20-MAY-1999; 99US-0135124.  
PR 21-MAY-1999; 99US-0135355.  
PR 24-MAY-1999; 99US-0135629.  
PR 25-MAY-1999; 99US-0136021.

PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 04-JUN-1999; 99US-0137502.  
PR 07-JUN-1999; 99US-0137724.  
PR 08-JUN-1999; 99US-0138094.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 17-JUN-1999; 99US-0139453.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139455.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139462.  
PR 18-JUN-1999; 99US-0139463.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139763.  
PR 21-JUN-1999; 99US-0139817.  
PR 22-JUN-1999; 99US-0139899.  
PR 23-JUN-1999; 99US-0140353.  
PR 23-JUN-1999; 99US-0140354.  
PR 24-JUN-1999; 99US-0140695.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141287.  
PR 01-JUL-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142055.  
PR 06-JUL-1999; 99US-0142390.  
PR 08-JUL-1999; 99US-0142803.  
PR 09-JUL-1999; 99US-0142977.  
PR 12-JUL-1999; 99US-0143297.  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143624.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144325.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 19-JUL-1999; 99US-0144333.  
PR 19-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144352.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144814.  
PR 21-JUL-1999; 99US-0145086.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145145.  
PR 23-JUL-1999; 99US-0145218.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145276.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145951.  
PR 28-JUL-1999; 99US-0145951.  
PR 02-AUG-1999; 99US-0146386.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146389.  
PR 03-AUG-1999; 99US-0147038.

```

PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147203.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151338.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156559.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160778.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 21-OCT-1999; 99US-0160880.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.

```

```

PR 29-OCT-1999; 99US-0162142.
Query Match 76.3%; Score 29; DB 21; Length 415;
Best Local Similarity 75.0%; Pred. No. 6; 7e-02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
OY 1 VLXDDLE 8
Db 66 LLIDDLLE 73
RESULT 42
AAG61477
ID AAG61477 standard; Protein; 415 AA.
XX
AC AAG61477;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 79744.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134321.
PR 14-MAY-1999; 99US-0134320.
PR 18-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.

```

PR 10-JUN-1999; 99US-0138540.  
 PR 10-JUN-1999; 99US-0138847.  
 PR 14-JUN-1999; 99US-0139119.  
 PR 16-JUN-1999; 99US-0139452.  
 PR 16-JUN-1999; 99US-0139453.  
 PR 17-JUN-1999; 99US-0139454.  
 PR 18-JUN-1999; 99US-0139455.  
 PR 18-JUN-1999; 99US-0139456.  
 PR 18-JUN-1999; 99US-0139457.  
 PR 18-JUN-1999; 99US-0139458.  
 PR 18-JUN-1999; 99US-0139459.  
 PR 18-JUN-1999; 99US-0139460.  
 PR 18-JUN-1999; 99US-0139461.  
 PR 18-JUN-1999; 99US-0139462.  
 PR 18-JUN-1999; 99US-0139463.  
 PR 18-JUN-1999; 99US-0139750.  
 PR 18-JUN-1999; 99US-0139763.  
 PR 21-JUN-1999; 99US-0139817.  
 PR 22-JUN-1999; 99US-0139899.  
 PR 23-JUN-1999; 99US-0140353.  
 PR 23-JUN-1999; 99US-0140354.  
 PR 24-JUN-1999; 99US-0140695.  
 PR 28-JUN-1999; 99US-0140823.  
 PR 29-JUN-1999; 99US-0140991.  
 PR 30-JUN-1999; 99US-0141287.  
 PR 01-JUL-1999; 99US-0141842.  
 PR 01-JUL-1999; 99US-0142154.  
 PR 02-JUL-1999; 99US-0142055.  
 PR 06-JUL-1999; 99US-0142390.  
 PR 08-JUL-1999; 99US-0142803.  
 PR 09-JUL-1999; 99US-0142920.  
 PR 12-JUL-1999; 99US-0143542.  
 PR 13-JUL-1999; 99US-0143624.  
 PR 14-JUL-1999; 99US-0144005.  
 PR 15-JUL-1999; 99US-0144086.  
 PR 16-JUL-1999; 99US-0144087.  
 PR 19-JUL-1999; 99US-0144332.  
 PR 19-JUL-1999; 99US-0144331.  
 PR 19-JUL-1999; 99US-0144332.  
 PR 19-JUL-1999; 99US-0144333.  
 PR 19-JUL-1999; 99US-0144334.  
 PR 19-JUL-1999; 99US-0144335.  
 PR 20-JUL-1999; 99US-0144352.  
 PR 20-JUL-1999; 99US-0144632.  
 PR 21-JUL-1999; 99US-0144884.  
 PR 21-JUL-1999; 99US-0144814.  
 PR 21-JUL-1999; 99US-0145086.  
 PR 21-JUL-1999; 99US-0145088.  
 PR 22-JUL-1999; 99US-0145087.  
 PR 22-JUL-1999; 99US-0145087.  
 PR 22-JUL-1999; 99US-0145087.  
 PR 22-JUL-1999; 99US-0145089.  
 PR 22-JUL-1999; 99US-0145192.  
 PR 22-JUL-1999; 99US-0145145.  
 PR 23-JUL-1999; 99US-0145218.  
 PR 23-JUL-1999; 99US-0145224.  
 PR 26-JUL-1999; 99US-0145276.  
 PR 27-JUL-1999; 99US-0145913.  
 PR 27-JUL-1999; 99US-0145918.  
 PR 27-JUL-1999; 99US-0145919.  
 PR 27-JUL-1999; 99US-0145951.  
 PR 28-JUL-1999; 99US-0146386.  
 PR 02-AUG-1999; 99US-0146389.  
 PR 02-AUG-1999; 99US-0146389.  
 PR 03-AUG-1999; 99US-0147038.  
 PR 04-AUG-1999; 99US-0147204.  
 PR 04-AUG-1999; 99US-0147302.  
 PR 05-AUG-1999; 99US-0147192.  
 PR 05-AUG-1999; 99US-0147260.  
 PR 06-AUG-1999; 99US-0147303.  
 PR 06-AUG-1999; 99US-0147416.  
 PR 09-AUG-1999; 99US-0147493.

PR 09-AUG-1999; 99US-0147935.  
 PR 10-AUG-1999; 99US-0148171.  
 PR 11-AUG-1999; 99US-0148319.  
 PR 12-AUG-1999; 99US-0148341.  
 PR 13-AUG-1999; 99US-0148565.  
 PR 13-AUG-1999; 99US-0148684.  
 PR 16-AUG-1999; 99US-0149368.  
 PR 16-AUG-1999; 99US-0149175.  
 PR 17-AUG-1999; 99US-0149426.  
 PR 18-AUG-1999; 99US-0149722.  
 PR 20-AUG-1999; 99US-0149723.  
 PR 20-AUG-1999; 99US-0149929.  
 PR 23-AUG-1999; 99US-0149902.  
 PR 23-AUG-1999; 99US-0149930.  
 PR 25-AUG-1999; 99US-0150566.  
 PR 25-AUG-1999; 99US-0150884.  
 PR 27-AUG-1999; 99US-0151065.  
 PR 27-AUG-1999; 99US-0151066.  
 PR 27-AUG-1999; 99US-0151080.  
 PR 30-AUG-1999; 99US-0151303.  
 PR 31-AUG-1999; 99US-0151438.  
 PR 01-SEP-1999; 99US-0151930.  
 PR 07-SEP-1999; 99US-0152363.  
 PR 10-SEP-1999; 99US-0153070.  
 PR 13-SEP-1999; 99US-0153758.  
 PR 15-SEP-1999; 99US-0154018.  
 PR 16-SEP-1999; 99US-0154039.  
 PR 20-SEP-1999; 99US-0154779.  
 PR 22-SEP-1999; 99US-0155139.  
 PR 23-SEP-1999; 99US-0155486.  
 PR 24-SEP-1999; 99US-0155659.  
 PR 28-SEP-1999; 99US-0156458.  
 PR 29-SEP-1999; 99US-0156596.  
 PR 04-OCT-1999; 99US-0157117.  
 PR 05-OCT-1999; 99US-0157753.  
 PR 06-OCT-1999; 99US-0157865.  
 PR 07-OCT-1999; 99US-0158029.  
 PR 08-OCT-1999; 99US-0158232.  
 PR 12-OCT-1999; 99US-0158369.  
 PR 13-OCT-1999; 99US-0159293.  
 PR 13-OCT-1999; 99US-0159294.  
 PR 13-OCT-1999; 99US-0159295.  
 PR 13-OCT-1999; 99US-0159329.  
 PR 14-OCT-1999; 99US-0159330.  
 PR 14-OCT-1999; 99US-0159331.  
 PR 14-OCT-1999; 99US-0159637.  
 PR 14-OCT-1999; 99US-0159637.  
 PR 14-OCT-1999; 99US-0159638.  
 PR 18-OCT-1999; 99US-0159584.  
 PR 21-OCT-1999; 99US-0160741.  
 PR 21-OCT-1999; 99US-0160767.  
 PR 21-OCT-1999; 99US-0160768.  
 PR 21-OCT-1999; 99US-0160770.  
 PR 21-OCT-1999; 99US-0160814.  
 PR 21-OCT-1999; 99US-0160815.  
 PR 22-OCT-1999; 99US-0160980.  
 PR 22-OCT-1999; 99US-0160981.  
 PR 22-OCT-1999; 99US-0161404.  
 PR 22-OCT-1999; 99US-0161405.  
 PR 25-OCT-1999; 99US-0161406.  
 PR 25-OCT-1999; 99US-0161359.  
 PR 26-OCT-1999; 99US-0161360.  
 PR 26-OCT-1999; 99US-0161361.  
 PR 26-OCT-1999; 99US-0161920.  
 PR 28-OCT-1999; 99US-0161923.  
 PR 28-OCT-1999; 99US-0161993.  
 PR 29-OCT-1999; 99US-0162142.

Query Match 76.3%; Score 29; DB 21; Length 415;  
 Best Local Similarity 75.0%; Pred. No. 6.7e+02;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 66 L1DLDLLE 73

## RESULT 43

ABB65976  
ID ABB65976 standard; Protein; 429 AA.

AC ABB65976;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 24720.

KW Drosophila: developmental biology; cell signalling; insecticide; pharmaceutical.

OS Drosophila melanogaster.

PN WO200171042-A2.

PD 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US09231.

PR 23-MAR-2000; 2000US-191637P.  
11-JUL-2000; 2000US-0614150.

PA (PEKE ) PE CORP NY.

PI Venter JC, Adams M, Li PWD, Myers EW.

DR WPI: 2001-656860/75.

DR N-PSDB: ABL10079.

PT New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions -

PS Disclosure: SEQ ID NO 24720; 21pp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL01840-ABL16175), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins (ABR57737-ABR72072).  
CC The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.

SQ Sequence 429 AA:

Query Match 76.3%; Score 29; DB 22; Length 429;  
Best Local Similarity 44.4%; Pred. No. 7e+02;

Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 1 VLXDDLLEA 9

DB 385 ILCDITFEA 393

## RESULT 44

ABB67967  
ID ABB67967 standard; Protein; 468 AA.

AC ABB67967;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 30693.

XX Drosophila: developmental biology; cell signalling; insecticide; pharmaceutical.

OS Drosophila melanogaster.

PN WO200171042-A2.

PD 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US09231.

PR 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

PA (PEKE ) PE CORP NY.

PI Venter JC, Adams M, Li PWD, Myers EW.

DR WPI: 2001-656860/75.

DR N-PSDB: ABL12070.

PT New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions -

PS Disclosure: SEQ ID NO 30693; 21pp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL01840-ABL16175), expressed DNA sequences (ABR57737-ABR72072).  
CC The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.

SQ Sequence 468 AA:

Query Match 76.3%; Score 29; DB 22; Length 468;  
Best Local Similarity 75.0%; Pred. No. 7.6e+02;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 VLXDDLLE 8

DB 92 VLXDDLLE 99

## RESULT 45

ABB67165  
ID ABB67165 standard; Protein; 505 AA.

AC ABB67165;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 28287.

KW Drosophila: developmental biology; cell signalling; insecticide; pharmaceutical.

OS Drosophila melanogaster.

PN WO200171042-A2.

PD 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US09231.

PR 23-MAR-2000; 2000US-191637P.







XX Homo sapiens.  
 OS  
 XX  
 PN EPI074617-A2.  
 XX  
 PD 07-FEB-2001.  
 XX  
 PF 28-JUL-2000; 2000EP-0116126.  
 XX  
 PR 29-JUL-1999; 99JP-0248036.  
 PR 27-AUG-1999; 99JP-0300253.  
 PR 11-JAN-2000; 2000JP-0118776.  
 PR 02-MAY-2000; 2000JP-0183767.  
 PR 09-JUN-2000; 2000JP-0241899.  
 XX  
 PA (HELI-) HELIX RES INST.  
 XX  
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
 PI Ishii S, Sugiyama T, Makamatsu A, Nagai K, Otsuki T;  
 XX  
 DR WPI: 2001-318749/34.  
 XX  
 PT Primer sets for synthesizing polynucleotides, particularly the 5602  
 PT full-length cDNAs defined in the specification, and for the detection  
 PT and/or diagnosis of the abnormality of the proteins encoded by the  
 PT full-length cDNAs -  
 XX  
 PS Claim 8; SEQ ID 18695; 2537bp + CD ROM; English.  
 XX  
 CC The present invention describes primer sets for synthesizing 5602  
 CC full-length cDNAs defined in the specification. Where a primer set  
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary  
 CC to the complementary strand of a polynucleotide which comprises one of  
 CC the 5602 nucleotide sequences defined in the specification, where the  
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
 CC of an oligonucleotide comprising a sequence complementary to the  
 CC complementary strand of a polynucleotide which comprises a 5'-end  
 CC sequence and an oligonucleotide comprising a sequence complementary to a  
 CC polynucleotide which comprises a 3'-end sequence, where the  
 CC oligonucleotide comprises at least 15 nucleotides and the combination of  
 CC the 5'-end sequence/3'-end sequence is selected from those defined in  
 CC the specification. The primer sets can be used in antisense therapy and  
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,  
 CC particularly full-length cDNAs. The primers are also useful for the  
 CC detection and/or diagnosis of the abnormality of the proteins encoded by  
 CC the full-length cDNAs. The primers allow obtaining of the full-length  
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to  
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632  
 CC represent oligonucleotides, all of which are used in the exemplification  
 CC of the present invention.  
 XX  
 SQ Sequence 559 AA;  
 XX  
 Query Match 76.3%; Score 29; DB 22; Length 559;  
 Best Local Similarity 66.7%; Pred. No. 9.2e+02;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 XX  
 QY 1 VLXDDLLEA 9  
 Db 307 ILDDLSEA 315  
 XX  
 RESULT 50  
 AAB10765  
 ID AAB10765 standard; Protein: 565 AA.  
 XX  
 AC AAB10765;  
 XX  
 DT 26-JAN-2001 (first entry)  
 XX  
 DE Murine T gene protein fragment.  
 XX

KW T gene: murine; central nervous system development; CNS; nootropic;  
 KW neuroleptic; antidepressant; gene therapy; antisense; treatment;  
 KW schizophrenia; autism; manic depression; mental retardation.  
 XX  
 OS Mus sp.  
 XX  
 PN DE1908423-A1.  
 XX  
 PD 31-AUG-2000.  
 XX  
 PR 26-FEB-1999; 99DE-1008423.  
 PR 26-FEB-1999; 99DE-1008423.  
 XX  
 PA (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.  
 XX  
 PI Poustka A, Coy J;  
 PI WPI: 2000-580150/55.  
 DR N-PSDB; AAA98007.  
 XX  
 PT DNA encoding a protein involved in development of the central nervous  
 PT system (CNS), antisense sequences, ribozymes and antibodies, useful for  
 PT treatment of, e.g. schizophrenia and manic depression -  
 XX  
 PS Claim 2; Fig 12; 86pp; German.  
 XX  
 CC This invention describes a novel DNA sequence, which encodes a protein  
 CC that is involved in development of the central nervous system (CNS) and  
 CC has tissue and development-specific expression. The products of the  
 CC invention have nootropic, neuroleptic and antidepressant activity and can  
 CC be used for gene therapy and antisense inhibition. The method also  
 CC describes a method for producing (1) antisense RNA that is complementary  
 CC to DNA as above, which can reduce or inhibit synthesis of the protein  
 CC coding DNA; (2) a ribozyme, which is complementary to DNA as above, which  
 CC specifically binds to and cleaves transcribed DNA, which reduces or  
 CC inhibits synthesis of the protein coding DNA; (3) an expression vector,  
 CC containing DNA as above, or which encodes antisense RNA or a ribozyme;  
 CC (4) a host cell transformed with a vector as in (3); (5) a protein,  
 CC encoded by DNA as above; (6) a method to produce the protein of (5)  
 CC comprising culturing the cell of (4) and isolating the protein from the  
 CC cell or the culture medium; (7) an antibody targeted against the protein  
 CC of (5); (8) a diagnostic method to detect disturbed expression of the  
 CC protein of (5) or to detect altered forms of the protein by contacting a  
 CC sample with a DNA sequence or antibody and determining direct or indirect  
 CC contact; and comparing the expression of the protein with a healthy  
 CC patient; (9) a diagnostic kit to perform the method of (8); (10) a  
 CC non-human transgenic animal, where the naturally occurring T gene has an  
 CC altered gene structure or sequence; and (11) a method to produce a  
 CC non-human animal as in (10). The DNA, derived from the T gene encodes a T  
 CC protein (rp) which is involved in development of the central nervous  
 CC system. Antisense sequences, ribozymes and antibodies are useful for  
 CC treatment of disorders of the CNS including schizophrenia, autism, manic  
 CC depression and mental retardation. This sequence represents a fragment of  
 CC the murine T protein described in the method of the invention.  
 XX  
 SQ Sequence 565 AA;  
 XX  
 Query Match 76.3%; Score 29; DB 21; Length 565;  
 Best Local Similarity 66.7%; Pred. No. 9.3e+02;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 XX  
 QY 1 VLXDDLLEA 9  
 Db 313 ILDDLSEA 321  
 XX  
 Search completed: December 27, 2002, 14:41:38  
 Job time : 57 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: December 27, 2002, 14:40:33 ; Search time 16 Seconds

(without alignments)  
54.076 Million cell updates/sec

Title: US-09-489-760a-1  
Perfect score: 38  
Sequence: 1 VLXDDLEA 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 1000 summaries

## Database :

1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	36	94.7	1165	2	D59433	C. elegans protein
2	34	89.5	616	2	C69226	type I restriction
3	33	86.8	174	2	C88075	protein R04A8.9 [i
4	33	86.8	499	2	G86538	leucyl aminopeptid
5	33	86.8	499	2	G72083	leucine aminopepti
6	32	84.2	112	2	S33180	nitrogen regulator
7	32	84.2	797	2	T50072	hypothetical prote
8	31	81.6	276	2	AB1663	sugar ABC transpor
9	31	81.6	276	2	AD1291	sugar ABC transpor
10	31	81.6	289	2	D81930	probable UTP-glucos
11	31	81.6	289	2	C81175	UTP-glucose-1-phos
12	31	81.6	289	2	A28180	carbamoyl-phosphat
13	31	81.6	497	2	G87793	protein C27A12.7 [
14	31	81.6	750	2	B83794	hypothetical prote
15	31	81.6	1042	2	T13647	hypothetical prote
16	30	78.9	72	2	T16975	hypothetical prote
17	30	78.9	302	2	AB3246	transposase tnp [l
18	30	78.9	361	2	T48014	serine/threonine p
19	30	78.9	623	2	E75221	type I restriction
20	30	78.9	683	2	D82674	TonB-dependent rec
21	30	78.9	822	2	I38728	epidermal growth f
22	30	78.9	962	1	UT0669	helicase II-like p
23	30	78.9	1999	2	AB2018	hypothetical prote
24	30	78.9	4976	2	T14165	peptide synthetase
25	29	76.3	78	2	C81976	hypothetical prote
26	29	76.3	101	2	E97582	hypothetical prote
27	29	76.3	101	2	AE2803	hypothetical prote
28	29	76.3	119	2	T35859	hypothetical prote
29	29	76.3	137	1	Q3ECEA	ydbb protein - Esc
30	29	76.3	137	2	D90708	hypothetical prote
31	29	76.3	137	2	H85558	hypothetical hypotet
32	29	76.3	137	2	A10575	conserved hypotet
33	29	76.3	154	1	R5HS30	ribosomal protein
34	29	76.3	156	2	T20267	hypothetical prote
35	29	76.3	178	2	D72293	hypothetical prote
36	29	76.3	192	2	A12271	transcription regu
37	29	76.3	207	2	G83538	probable oxidoredu
38	29	76.3	231	2	B65140	hypothetical 26.3
39	29	76.3	231	2	H91164	hypothetical prote
40	29	76.3	231	2	H86010	hypothetical prote
41	29	76.3	241	2	D86725	hypothetical prote
42	29	76.3	251	2	S45438	cytochrome-c oxida
43	29	76.3	277	2	D69158	sensory transduct
44	29	76.3	292	2	S72323	ubiquinol-cytochro
45	29	76.3	322	2	S44847	K06H7.7 protein -
46	29	76.3	323	2	E90949	heat shock protein
47	29	76.3	323	2	A85798	suppressor of htrB
48	29	76.3	323	2	A42608	(Kdo)2-(Jauray)-1
49	29	76.3	358	2	AC2367	glucose-1-phosphat
50	29	76.3	392	1	B69852	probable NADH2 deh
51	29	76.3	415	2	D95248	conserved hypotet
52	29	76.3	415	2	A98113	melibiase (alpha-g
53	29	76.3	434	2	D83528	conserved hypotet
54	29	76.3	440	1	B70326	protein F12M16.8 [
55	29	76.3	457	2	E96572	hypothetical prote
56	29	76.3	493	2	AB3032	hypothetical prote
57	29	76.3	495	2	T27693	beta-amylase (EC 3
58	29	76.3	496	2	T08117	phosphoribosylamin
59	29	76.3	530	2	C82344	GTP-binding membra
60	29	76.3	559	2	H70110	hypothetical prote
61	29	76.3	571	2	H87600	exse protein (Auj2
62	29	76.3	602	2	A96254	lesits nuclear RNA
63	29	76.3	619	2	I48840	probable transcrip
64	29	76.3	715	2	T36309	probable phosphopr
65	29	76.3	738	2	B84670	hypothetical prote
66	29	76.3	1002	2	AF2383	probable phosphopr
67	29	76.3	1088	2	E86217	hypothetical prote
68	29	76.3	1148	2	S51855	cleavage and polya
69	29	76.3	1444	2	S57335	ferredoxin [4Fe-4S
70	29	73.7	81	1	FERSFE	ferredoxin [4Fe-4S
71	28	73.7	81	1	A55790	hypothetical prote
72	28	73.7	84	2	E72497	nickel-responsive
73	28	73.7	133	2	S47700	nickel-responsive
74	28	73.7	133	2	D91172	nickel-responsive
75	28	73.7	133	2	D86018	nickel-responsive
76	28	73.7	133	2	AB0990	nickel-responsive
77	28	73.7	135	2	D44503	p19 protein - beet
78	28	73.7	137	2	D83748	hypothetical prote
79	28	73.7	163	2	F81339	probable membrane
80	28	73.7	178	2	AE3262	hypoxanthine phosph
81	28	73.7	180	2	G98283	hypoxanthine phosph
82	28	73.7	180	2	AB3000	BAG-family molecu
83	28	73.7	195	2	T39603	phosphatidylglycer
84	28	73.7	199	2	A87845	conserved hypotet
85	28	73.7	209	2	H82174	hypothetical prote
86	28	73.7	213	2	T14945	hypothetical prote
87	28	73.7	245	2	A84289	hypothetical prote
88	28	73.7	249	2	G69678	involved in polyke
89	28	73.7	252	1	MRIVC	matrix protein M1
90	28	73.7	252	1	B45539	matrix protein M1
91	28	73.7	252	1	MRIVS	matrix protein M1
92	28	73.7	252	1	MRIVK	matrix protein M1
93	28	73.7	252	1	MRIVL	matrix protein M1
94	28	73.7	252	1	MRIVM	matrix protein M1
95	28	73.7	252	1	PN0083	matrix protein M1
96	28	73.7	252	1	PN0086	matrix protein M1
97	28	73.7	252	1	S04054	matrix protein M1
98	28	73.7	252	2	S04052	matrix protein M1
99	28	73.7	252	2	S04058	matrix protein M1
100	28	73.7	252	2	S04050	matrix protein M1
101	28	73.7	252	2	S14616	matrix protein M1
102	28	73.7	252	2	S04056	matrix protein M1

103	28	73.7	257	2	F87209	probable enoyl-CoA
104	28	73.7	282	2	S31258	probable membrane
105	28	73.7	282	2	JC5677	RNA4 protein - Bee
106	28	73.7	282	2	C44503	p31 protein - bee
107	28	73.7	288	2	G81129	conserved hypotnet
108	28	73.7	288	2	D81837	hypothetical prote
109	28	73.7	295	2	C82558	bifunctional methyl
110	28	73.7	296	2	G95108	DNA polymerase III
111	28	73.7	296	2	D97976	DNA-directed DNA p
112	28	73.7	302	2	AE3127	transcription regu
113	28	73.7	302	2	D98160	oxidative stress t
114	28	73.7	314	2	B98232	Probable 2-hydroxy
115	28	73.7	327	2	AC3054	2-hydroxyacid dehy
116	28	73.7	343	2	A64443	glyceralddehyde-3-P
117	28	73.7	383	2	E75589	hypothetical prote
118	28	73.7	389	2	T20604	hypothetical prote
119	28	73.7	397	2	F90114	hypothetical prote
120	28	73.7	398	2	AB0712	26S protease regul
121	28	73.7	411	2	T47951	probable regulator
122	28	73.7	421	2	A99409	hypothetical prote
123	28	73.7	442	2	G83672	conserved hypotnet
124	28	73.7	486	2	A82878	6-phospho-beta-glu
125	28	73.7	517	2	S12327	replicative DNA he
126	28	73.7	522	2	C90073	hypothetical prote
127	28	73.7	538	2	S76175	hypothetical prote
128	28	73.7	543	2	B84398	hypothetical prote
129	28	73.7	550	1	A49936	arginine-tRNA liga
130	28	73.7	558	2	S61604	probable membrane
131	28	73.7	612	2	T39666	WD-repeat protein
132	28	73.7	776	2	AH1147	preprotein translo
133	28	73.7	776	2	AH1506	preprotein translo
134	28	73.7	837	2	T12514	hypothetical prote
135	28	73.7	867	2	F81086	aminopeptidase N N
136	28	73.7	867	2	G81856	membrane alanyl am
137	28	73.7	910	2	S38167	hypothetical prote
138	28	73.7	980	2	E71606	hypothetical prote
139	28	73.7	1104	2	B75221	type I restriction
140	28	73.7	1231	2	C84716	hypothetical prote
141	28	73.7	1277	2	F81132	oxidoreductase, pr
142	28	73.7	1277	2	G81888	probable oxidoredu
143	28	73.7	1408	2	T45039	hypothetical prote
144	28	73.7	1420	2	T17158	CL2AB protein - ra
145	28	73.7	1435	2	T46611	CL2BA protein - ra
146	28	73.7	1452	2	T17157	CL2BA protein - ra
147	28	73.7	1463	2	T17159	CL2AC protein - ra
148	28	73.7	1467	2	T17160	CL2BA protein - ra
149	28	73.7	1478	2	T17185	CL2BC protein - ra
150	28	73.7	1487	2	T14324	alpha-lactotoxin r
151	28	73.7	2244	2	F90563	hypothetical prote
152	28	73.7	3519	2	S43048	polyketide synthas
153	27	71.1	111	2	A33600	nitrogen regulator
154	27	71.1	116	1	R5HS86	ribosomal protein
155	27	71.1	119	2	T00151	hypothetical prote
156	27	71.1	155	1	E64168	anaerobic ribonuc
157	27	71.1	155	1	S23357	H+-transporing tw
158	27	71.1	158	2	T02973	probenaazole-induc
159	27	71.1	170	2	F97403	polypeptide deform
160	27	71.1	170	2	AF2621	polypeptide deform
161	27	71.1	172	2	B81057	conserved hypotnet
162	27	71.1	172	2	H81820	probable lipoprote
163	27	71.1	179	2	D64591	adenine phosphorib
164	27	71.1	179	2	H71920	cGMP-gated cation
165	27	71.1	189	2	S74159	cGMP-gated cation
166	27	71.1	191	2	S74158	hypothetical dihyd
167	27	71.1	191	2	AF1787	hypothetical prote
168	27	71.1	198	2	T28191	phosphatidylglycer
169	27	71.1	198	2	AE8435	hypothetical prote
170	27	71.1	201	2	B71635	conserved hypotnet
171	27	71.1	201	2	C87510	hypothetical prote
172	27	71.1	206	2	D72323	hypothetical prote
173	27	71.1	213	2	E71190	hypothetical prote
174	27	71.1	218	2	A75150	hypothetical prote
175	27	71.1	221	2	AF3379	molybdopterin-guan
176	27	71.1	239	2	S39723	spore coat polysac
177	27	71.1	244	2	D82123	RNA polymerase sig
178	27	71.1	248	2	G83238	hypothetical prote
179	27	71.1	256	2	AB2958	transcription regu
180	27	71.1	261	2	T35708	glutathione - Strept
181	27	71.1	263	2	D95180	glutathione S-tran
182	27	71.1	263	2	H98047	conserved hypotnet
183	27	71.1	264	2	T19511	hypothetical prote
184	27	71.1	275	2	E91102	probable enzyme (i
185	27	71.1	275	2	A85948	probable enzyme y9
186	27	71.1	275	2	F65076	hypothetical prote
187	27	71.1	279	2	T48106	hypothetical prote
188	27	71.1	281	2	AT0461	sn glycerol-3-phos
189	27	71.1	283	2	D98325	hypothetical prote
190	27	71.1	287	2	AG3118	ABC transporter, m
191	27	71.1	288	2	H98168	hypothetical prote
192	27	71.1	289	2	F47070	CDP-abequose synth
193	27	71.1	295	2	C75054	methionyl aminopep
194	27	71.1	297	2	D67187	phosphoribosylamin
195	27	71.1	299	2	D75138	hypothetical prote
196	27	71.1	300	2	DB3714	surface adhesin A
197	27	71.1	309	1	CCBY1H	ubiquinol-cytochro
198	27	71.1	310	2	C96929	ABC-type multidrug
199	27	71.1	311	2	AB3218	probable 2-hydroxy
200	27	71.1	311	2	D75365	UDP-glucose 4-epim
201	27	71.1	311	2	H90156	probable H+-transp
202	27	71.1	323	2	T01103	quinone oxidoreduc
203	27	71.1	324	2	A86839	conserved hypotnet
204	27	71.1	324	2	T18790	hypothetical prote
205	27	71.1	324	2	B90805	probable dehydroge
206	27	71.1	325	2	G85664	probable dehydroge
207	27	71.1	325	2	E64845	probable 2-hydroxy
208	27	71.1	325	2	F84740	hypothetical prote
209	27	71.1	329	1	DMECMD	threonine ammonia-
210	27	71.1	329	2	D85973	threonine ammonia-
211	27	71.1	329	2	E91128	hypothetical prote
212	27	71.1	329	2	T24285	site-specific DNA-
213	27	71.1	331	2	DM0649	protoporphyrin IX
214	27	71.1	335	2	C95118	Cdc6 related prote
215	27	71.1	335	2	A97988	hypothetical prote
216	27	71.1	358	2	S15908	hypothetical prote
217	27	71.1	358	2	G64413	probable arylsulfa
218	27	71.1	365	2	D69080	branched-chain-aml
219	27	71.1	379	1	G83808	gamma-glutamyl pho
220	27	71.1	382	2	G72225	protein T07E3.4 (I
221	27	71.1	385	2	E71238	fixc protein homol
222	27	71.1	385	2	E71238	flavoprotein, prob
223	27	71.1	395	2	A86286	xre family DNA-bin
224	27	71.1	401	2	A47014	probable phosphori
225	27	71.1	410	2	T46034	UDP-N-acetylglucos
226	27	71.1	421	2	AD2305	hypothetical prote
227	27	71.1	427	2	G88492	hypothetical prote
228	27	71.1	429	2	C64928	siroheme synthase
229	27	71.1	429	2	F90929	transcription regu
230	27	71.1	434	2	B85778	probable transposo
231	27	71.1	434	2	B97018	siroheme synthase
232	27	71.1	435	2	H71138	protein F25E5.11 (
233	27	71.1	445	2	E72229	replication initia
234	27	71.1	449	2	G89841	D-stereospecific a
235	27	71.1	454	2	C70147	replication initia
236	27	71.1	459	2	T48598	N2,N2-dimethylguan
237	27	71.1	464	2	T28662	probable phosphori
238	27	71.1	475	2	AE3036	trRNA (guanine-N2-)
239	27	71.1	481	2	AD1337	
240	27	71.1	483	2	H85073	
241	27	71.1	488	2	F98249	
242	27	71.1	489	2	E89102	
243	27	71.1	513	1	ZABPE4	
244	27	71.1	520	1	A42209	
245	27	71.1	522	2	JS0450	
246	27	71.1	524	2	T39993	
247	27	71.1	545	2	T00485	
248	27	71.1	548	2	T46565	

249	27	71.1	550	2	A87050	322	26	68.4	148	2	T11572	lysosome (EC 3.2.1
250	27	71.1	570	2	T27407	323	26	68.4	164	2	T21240	hypothetical prote
251	27	71.1	578	2	C64452	324	26	68.4	167	2	G90175	NADH dehydrogenase
252	27	71.1	602	2	T01360	325	26	68.4	169	2	F98272	hypothetical prote
253	27	71.1	603	2	S76959	326	26	68.4	169	2	A13011	conserved hypotnet
254	27	71.1	605	2	AG0123	327	26	68.4	173	2	G87282	polypeptide deform
255	27	71.1	617	2	C72670	328	26	68.4	173	2	T21815	hypothetical prote
256	27	71.1	626	2	AE0123	329	26	68.4	174	2	S39877	carc protein - Myx
257	27	71.1	628	2	T39623	330	26	68.4	178	2	E82900	hypothetical prote
258	27	71.1	645	2	T50680	331	26	68.4	178	2	A95217	transcription regu
259	27	71.1	663	2	S11521	332	26	68.4	182	2	G98080	conserved hypotnet
260	27	71.1	664	2	S11517	333	26	68.4	182	2	B75089	probable orotate p
261	27	71.1	682	1	JH0560	334	26	68.4	185	2	T47129	hypoxanthine-guanl
262	27	71.1	683	2	T51578	335	26	68.4	185	2	A82569	hypothetical prote
263	27	71.1	685	2	A81246	336	26	68.4	186	2	C95207	hypothetical prote
264	27	71.1	685	2	G82022	337	26	68.4	191	2	T35792	beta-glucosidase -
265	27	71.1	686	1	A44842	338	26	68.4	200	2	H70384	hypothetical prote
266	27	71.1	687	2	T09994	339	26	68.4	203	2	AC3611	hypothetical prote
267	27	71.1	690	1	S07103	340	26	68.4	208	2	S73861	hypothetical prote
268	27	71.1	690	2	A42161	341	26	68.4	210	2	G96832	protein associated
269	27	71.1	691	2	JC6509	342	26	68.4	210	2	G83929	conserved hypotnet
270	27	71.1	695	2	S74179	343	26	68.4	219	2	AB3498	probable orotate p
271	27	71.1	695	2	A45985	344	26	68.4	226	2	C87637	orotate phosphorib
272	27	71.1	704	2	A43411	345	26	68.4	226	2	J00928	hypoxanthine-guanl
273	27	71.1	706	2	T26218	346	26	68.4	229	2	T28284	hypothetical prote
274	27	71.1	707	2	S35691	347	26	68.4	230	2	A10994	hypothetical prote
275	27	71.1	732	2	I50630	348	26	68.4	231	2	D70032	conserved hypotnet
276	27	71.1	735	2	S39983	349	26	68.4	237	2	S71022	coat protein - tom
277	27	71.1	821	2	S73939	350	26	68.4	237	2	G64236	hypothetical prote
278	27	71.1	851	2	G83635	351	26	68.4	239	2	S11895	conserved hypotnet
279	27	71.1	902	2	D89057	352	26	68.4	239	2	JC4346	transcription init
280	27	71.1	974	2	H81311	353	26	68.4	239	2	E50961	transcription init
281	27	71.1	978	2	DUBB11	354	26	68.4	239	2	E85809	RNA polymerase sig
282	27	71.1	985	1	S65668	355	26	68.4	239	2	AF0750	RNA polymerase sig
283	27	71.1	1011	2	B85023	356	26	68.4	240	2	AH0224	hypothetical prote
284	27	71.1	1036	2	A57386	357	26	68.4	240	2	A92225	minor tail protein
285	27	71.1	1036	2	E70760	358	26	68.4	243	2	T29727	hypothetical prote
286	27	71.1	1041	2	B27672	359	26	68.4	251	2	T13104	hypothetical prote
287	27	71.1	1051	2	E87058	360	26	68.4	251	2	DA4288	urea protein homol
288	27	71.1	1059	2	T00330	361	26	68.4	261	2	T41578	hypothetical prote
289	27	71.1	1081	2	T35822	362	26	68.4	261	2	A64075	probable DNA-(apur
290	27	71.1	1091	2	T33153	363	26	68.4	264	2	C72770	probable enoyl-CoA
291	27	71.1	1101	2	S26358	364	26	68.4	264	2	B83395	phage-related prot
292	27	71.1	1116	2	T15890	365	26	68.4	265	2	E84092	hypothetical prote
293	27	71.1	1168	2	F70535	366	26	68.4	267	2	AC3125	nonstructural prot
294	27	71.1	1172	2	F72233	367	26	68.4	277	2	G87775	hydrogenase-1 cofa
295	27	71.1	1206	2	S63247	368	26	68.4	283	1	J01637	hydrogenase-1 cofa
296	27	71.1	1238	2	T04456	369	26	68.4	283	2	J02324	hypothetical 32.5k
297	27	71.1	1548	2	BVFEFL	370	26	68.4	284	1	S35660	hypothetical prote
298	27	71.1	1597	1	T08428	371	26	68.4	286	2	J01548	hypothetical prote
299	27	71.1	1615	1	WMTM18	372	26	68.4	289	2	AB3656	chitinase (EC 3.2.
300	27	71.1	1615	1	WMTM18	373	26	68.4	292	2	T10106	histone deacetylase
301	27	71.1	1615	1	WMTM18	374	26	68.4	292	2	F59443	conserved hypotnet
302	27	71.1	1651	1	E87058	375	26	68.4	296	2	E87134	conserved hypotnet
303	27	71.1	1722	1	MYBYK	376	26	68.4	298	2	F97685	geranyltransferase
304	27	71.1	1839	1	S12332	377	26	68.4	304	2	A13285	conserved hypotnet
305	27	71.1	2194	1	GNNME7	378	26	68.4	304	2	AD2601	deoxyhypusine synt
306	27	71.1	2712	2	T05113	379	26	68.4	305	2	E69036	hypothetical prote
307	27	71.1	5149	2	F83545	380	26	68.4	307	2	D64752	conserved hypotnet
308	27	71.1	5825	2	T12117	381	26	68.4	309	2	H87566	phosphoribosyl pyr
309	27	71.1	73	2	A42508	382	26	68.4	311	2	E82532	phosphatidate cytl
310	27	71.1	73	2	T30790	383	26	68.4	312	2	G70114	E1B 19K/Bcl-2-inte
311	27	71.1	73	2	G97135	384	26	68.4	312	2	T51280	hypothetical prote
312	27	71.1	77	2	AG1358	385	26	68.4	314	2	S18674	hypothetical prote
313	27	71.1	77	2	AG1358	386	26	68.4	319	2	T51280	hypothetical prote
314	27	71.1	82	2	AF1856	387	26	68.4	325	2	S18575	hypothetical prote
315	27	71.1	85	2	D64367	388	26	68.4	325	2	C89844	hypothetical prote
316	27	71.1	106	2	AD1063	389	26	68.4	326	2	A44505	hypothetical prote
317	27	71.1	118	2	E69047	390	26	68.4	326	2	F95319	hypothetical prote
318	27	71.1	130	2	AB1316	391	26	68.4	327	2	AC0285	hypothetical prote
319	27	71.1	136	2	C70761	392	26	68.4	332	2	T35917	hypothetical prote
320	27	71.1	138	2	AG0293	393	26	68.4				
321	27	71.1	146	2	G75544	394	26	68.4				

395	26	68.4	335	2	AH2910	geranyltransferansf	468	26	68.4	485	2	G84961	cardiolipin synthe
396	26	68.4	335	2	A64588	hypothetical 37.8	469	26	68.4	494	2	G69433	signal-transducin
397	26	68.4	335	2	F91013	nucleoid-associated	470	26	68.4	495	2	G84193	Glu-tRNA amidotran
398	26	68.4	335	2	AC0786	nucleoid-associated	471	26	68.4	496	1	A60473	beta-amylase (EC 3
399	26	68.4	335	2	H85857	nucleoid-associated	472	26	68.4	498	1	S32859	oute protein - Erw
400	26	68.4	335	2	T24185	hypothetical prote	473	26	68.4	499	1	JC1447	beta-amylase (EC 3
401	26	68.4	341	2	T40365	probable ribose-ph	474	26	68.4	503	2	A12342	hypothetical prote
402	26	68.4	342	2	T06272	farnesyl-pyrophosp	475	26	68.4	503	2	AC2446	hypothetical prote
403	26	68.4	342	2	T26577	hypothetical prote	476	26	68.4	505	2	T37975	probable alanine a
404	26	68.4	344	2	T26577	conserved hypothet	477	26	68.4	509	2	T37975	conserved hypothet
405	26	68.4	345	2	B90176	probable alcohol d	478	26	68.4	513	2	F82216	transcription regu
406	26	68.4	346	1	H69789	probable alcohol d	479	26	68.4	513	2	T14864	probable monosach
407	26	68.4	347	2	T02669	hypothetical prote	480	26	68.4	515	2	F96829	hypothetical prote
408	26	68.4	348	2	F97453	lytB protein (Ae00	481	26	68.4	525	2	A87468	malate synthase li
409	26	68.4	348	2	AH2671	penicillin toleran	482	26	68.4	525	2	E98332	hypothetical prote
410	26	68.4	354	2	B75555	hypothetical prote	483	26	68.4	525	2	AR2950	GdpP family prote
411	26	68.4	355	2	T45995	hypothetical prote	484	26	68.4	525	2	AR2950	cell fusion glycop
412	26	68.4	356	2	D96537	hypothetical prote	485	26	68.4	546	2	T08210	probable args prot
413	26	68.4	357	2	T06746	hypothetical prote	486	26	68.4	550	2	H70772	hypothetical prote
414	26	68.4	357	2	S20736	hypothetical prote	487	26	68.4	558	2	H70772	hypothetical prote
415	26	68.4	357	1	A47620	porin precursor -	488	26	68.4	587	2	D83182	hypothetical prote
416	26	68.4	364	2	H70776	3-isopropylmalate	489	26	68.4	591	2	A69159	protoporphyrin IX
417	26	68.4	368	2	T21024	hypothetical prote	490	26	68.4	598	2	S51386	hypothetical prote
418	26	68.4	371	2	E97383	hypothetical prote	491	26	68.4	599	1	A34231	sulfite reductase
419	26	68.4	372	2	D70753	hypothetical prote	492	26	68.4	599	1	H65057	sulfite reductase
420	26	68.4	373	2	F84396	probable oxidoredu	493	26	68.4	599	1	C91081	sulfite reductase
421	26	68.4	379	2	E75364	signal recognition	494	26	68.4	599	2	D85926	sulfite reductase
422	26	68.4	379	2	G75428	conserved hypothet	495	26	68.4	599	2	AG0858	sulfite reductase
423	26	68.4	381	2	H83985	conserved hypothet	496	26	68.4	604	2	S36487	El protein - human
424	26	68.4	389	2	T34766	alpha-D-mannose-al	497	26	68.4	604	2	T08302	hypothetical prote
425	26	68.4	397	2	S67061	iron-sulfur cofact	498	26	68.4	604	2	T08222	hypothetical prote
426	26	68.4	397	2	E69173	hypothetical prote	499	26	68.4	605	2	S36459	El protein
427	26	68.4	400	2	A47094	protoporphyrin IX	500	26	68.4	606	2	AE0409	sulfite reductase
428	26	68.4	400	2	E87439	aspartate transami	501	26	68.4	610	2	E84513	probable DNA repli
429	26	68.4	400	2	C97623	aspartate aminotra	502	26	68.4	627	2	D71729	penicillin-like protei
430	26	68.4	400	2	AC2846	aspartate aminotra	503	26	68.4	627	2	A97729	daak-c-type molecula
431	26	68.4	401	1	G64189	acetyl-CoA	504	26	68.4	630	2	D87739	protein T28F2.7 (I
432	26	68.4	407	2	S66260	metallopeptinase	505	26	68.4	630	2	T15144	hypothetical prote
433	26	68.4	408	2	AH3269	dihydroliipoamide S	506	26	68.4	631	2	G70188	transcription init
434	26	68.4	408	2	F85023	probable transposia	507	26	68.4	636	2	A84252	transcription init
435	26	68.4	409	2	A70563	probable transposia	508	26	68.4	641	2	C84726	probable acyl-CoA
436	26	68.4	411	2	B83782	flavohemoglobin hm	509	26	68.4	643	2	T04308	probable acyl-CoA
437	26	68.4	413	2	S60930	probable membrane	510	26	68.4	660	2	T03038	DNA helicase (EC 3
438	26	68.4	417	2	E83333	hypothetical prote	511	26	68.4	666	1	D69103	mismatch repair pr
439	26	68.4	422	2	B84554	probable phospho	512	26	68.4	669	2	F84377	hypothetical prote
440	26	68.4	428	2	B89606	protein B0416.6 (I	513	26	68.4	675	2	T21027	translation elonga
441	26	68.4	431	2	B8162	hypothetical prote	514	26	68.4	682	2	B84415	hypothetical prote
442	26	68.4	432	2	H97580	hypothetical prote	515	26	68.4	684	2	A52004	hypothetical prote
443	26	68.4	432	2	AF2801	hypothetical prote	516	26	68.4	686	2	A52004	hypothetical prote
444	26	68.4	434	2	B90695	inosine-guanosine	517	26	68.4	686	2	H90226	multichromosome mal
445	26	68.4	434	2	F85545	inosine-guanosine	518	26	68.4	686	2	A51495	transcription anti
446	26	68.4	434	2	J00812	inosine kinase (EC	519	26	68.4	687	2	AF1137	transcription anti
447	26	68.4	438	2	AB1140	6-phospho-beta-glu	520	26	68.4	687	2	T34082	conserved hypothet
448	26	68.4	440	2	F91206	probable 6-phospho	521	26	68.4	690	2	H82923	hypothetical prote
449	26	68.4	442	2	B86052	hypothetical prote	522	26	68.4	696	2	A98155	hypothetical prote
450	26	68.4	443	2	T39457	hypothetical prote	523	26	68.4	696	2	B66000	hypothetical prote
451	26	68.4	447	2	T10876	Y4Kv protein - Rhi	524	26	68.4	696	2	B66000	hypothetical prote
452	26	68.4	448	2	S24756	vicillin-like stora	525	26	68.4	700	2	T27363	hypothetical prote
453	26	68.4	449	2	T43340	alpha-1,3-mannosyl	526	26	68.4	704	2	G83950	hypothetical prote
454	26	68.4	449	2	G97813	hypothetical prote	527	26	68.4	706	2	T36176	polynucleotide pho
455	26	68.4	452	2	T40769	hypothetical prote	528	26	68.4	709	2	CH0126	hypothetical prote
456	26	68.4	455	2	B86905	GMP-binding protei	529	26	68.4	711	2	C83922	mba protein precu
457	26	68.4	457	2	H71028	hypothetical prote	530	26	68.4	712	2	A47718	reverse transcripti
458	26	68.4	457	2	B96037	hypothetical prote	531	26	68.4	716	2	AC2449	ABC transporter AM
459	26	68.4	457	2	D95194	probable alpha-gal	532	26	68.4	726	2	E83712	chemotaxis protein
460	26	68.4	459	2	F97125	UDP-N-acetylglucam	533	26	68.4	728	1	S07558	ribonucleoside-dip
461	26	68.4	460	2	T27759	tlid protein trunc	534	26	68.4	730	2	A26391	translation elonga
462	26	68.4	463	2	T27411	hypothetical prote	535	26	68.4	730	2	B83540	phospholipase C (B
463	26	68.4	463	2	T27411	hypothetical prote	536	26	68.4	732	2	F84394	hemolytic phosphol
464	26	68.4	468	2	A37176	glutamate-ammonia	537	26	68.4	735	2	T01956	regulatory protein
465	26	68.4	478	2	D30169	leukotoxin secreti	538	26	68.4	735	2	S74209	multifunctional be
466	26	68.4	479	2	T23508	hypothetical prote	539	26	68.4	736	2	S59136	estradiol 17beta-d
467	26	68.4	480	2	T27413	hypothetical prote	540	26	68.4	739	2	T10932	polyribonucleotide

541	26	68.4	742	2	I37225	614	25	65.8	73	2	T28476	hypothetical prote
542	26	68.4	743	2	D75590	615	25	65.8	73	2	D72155	E14f protein - var
543	26	68.4	758	1	A39343	616	25	65.8	76	2	AG1590	hypothetical prote
544	26	68.4	765	2	E96558	617	25	65.8	80	2	H86967	hypothetical prote
545	26	68.4	769	2	C83710	618	25	65.8	82	2	D69621	hypothetical prote
546	26	68.4	770	2	T00203	619	25	65.8	86	2	T44923	hypothetical prote
547	26	68.4	776	2	D98354	620	25	65.8	89	2	T08295	hypothetical prote
548	26	68.4	776	2	AC2928	621	25	65.8	95	2	S49553	conserved hypotet
549	26	68.4	791	2	H72258	622	25	65.8	102	2	F69475	hypothetical prote
550	26	68.4	803	2	F59433	623	25	65.8	105	2	B70578	hypothetical prote
551	26	68.4	807	2	A86740	624	25	65.8	110	2	T00142	hypothetical prote
552	26	68.4	810	1	A33380	625	25	65.8	112	2	H81961	nitrogen regulator
553	26	68.4	814	2	A95206	626	25	65.8	112	2	H81019	hypothetical prote
554	26	68.4	818	2	A59433	627	25	65.8	113	2	G69065	disease resistance
555	26	68.4	822	2	T47007	628	25	65.8	120	2	T07765	cell division prot
556	26	68.4	822	2	AB0238	629	25	65.8	121	2	GA5278	cell division prot
557	26	68.4	826	2	T46060	630	25	65.8	121	2	AF0517	cell division prot
558	26	68.4	826	2	T46061	631	25	65.8	121	2	G85490	hypothetical prote
559	26	68.4	829	2	E64114	632	25	65.8	123	2	T32592	Chey homolog DivK
560	26	68.4	844	2	T52396	633	25	65.8	123	2	S58671	conserved hypotet
561	26	68.4	846	2	T38840	634	25	65.8	125	2	R87554	hypothetical prote
562	26	68.4	847	2	G95843	635	25	65.8	130	2	AF0703	adenine phosphorib
563	26	68.4	857	2	T20318	636	25	65.8	130	2	B69947	two-component syst
564	26	68.4	863	2	B72344	637	25	65.8	131	2	S73769	hypothetical prote
565	26	68.4	884	2	S53396	638	25	65.8	133	2	A11861	hypothetical prote
566	26	68.4	889	2	T02240	639	25	65.8	134	2	B83529	hypothetical prote
567	26	68.4	935	2	S57080	640	25	65.8	136	2	AF0928	conserved hypotet
568	26	68.4	937	2	B86210	641	25	65.8	136	2	E85776	hypothetical prote
569	26	68.4	975	2	T03004	642	25	65.8	136	2	H81020	conserved hypotet
570	26	68.4	979	2	T08316	643	25	65.8	139	2	H81963	hypothetical prote
571	26	68.4	990	2	H86293	644	25	65.8	139	2	B69388	hypothetical prote
572	26	68.4	995	2	AF2211	645	25	65.8	146	2	S49617	conserved hypotet
573	26	68.4	1042	2	H75112	646	25	65.8	150	2	T42690	hypothetical prote
574	26	68.4	1043	2	S38034	647	25	65.8	150	2	AF6317	conserved hypotet
575	26	68.4	1050	2	A89769	648	25	65.8	150	2	E81672	hypothetical prote
576	26	68.4	1063	2	S18211	649	25	65.8	150	2	E96521	LSU ribosomal prot
577	26	68.4	1153	2	F84468	650	25	65.8	154	2	T07760	protein F2108.14
578	26	68.4	1162	2	D83454	651	25	65.8	160	1	WZBEC7	disease resistance
579	26	68.4	1171	1	F83110	652	25	65.8	165	2	B87672	gene 34 protein -
580	26	68.4	1174	1	H18YDH	653	25	65.8	168	2	C82494	acetyltransferase,
581	26	68.4	1183	2	F90559	654	25	65.8	169	2	T09904	polypeptide deform
582	26	68.4	1205	2	D83862	655	25	65.8	169	2	A09584	hypothetical prote
583	26	68.4	1223	2	S43579	656	25	65.8	178	2	E81672	hypothetical prote
584	26	68.4	1255	2	T06916	657	25	65.8	178	2	AG0536	conserved hypotet
585	26	68.4	1272	2	C64513	658	25	65.8	180	1	RTMSA	adenine phosphorib
586	26	68.4	1333	2	S63403	659	25	65.8	180	2	A96993	macrophage infecti
587	26	68.4	1355	2	T00961	660	25	65.8	182	2	C95047	probable adenine p
588	26	68.4	1387	2	JC5502	661	25	65.8	182	2	AF3311	gas-vesicle operon
589	26	68.4	1468	1	S30818	662	25	65.8	186	2	T38223	hypothetical prote
590	26	68.4	1505	2	S28079	663	25	65.8	186	2	T08240	hypothetical prote
591	26	68.4	1514	2	T52080	664	25	65.8	191	2	S15184	hypothetical prote
592	26	68.4	1520	2	T00273	665	25	65.8	195	2	H95254	hypothetical prote
593	26	68.4	1592	2	S48933	666	25	65.8	195	2	G98119	phosphoglycerate m
594	26	68.4	1607	2	T04583	667	25	65.8	196	2	C70591	SSU ribosomal prot
595	26	68.4	1639	2	T14181	668	25	65.8	196	2	AI3440	probable transcrip
596	26	68.4	1661	2	S64800	669	25	65.8	198	2	H69487	nus operon 15k pro
597	26	68.4	1777	2	AC2088	670	25	65.8	198	2	T34578	conserved hypotet
598	26	68.4	1956	2	T16416	671	25	65.8	198	2	E97369	thermophilic NADP
599	26	68.4	1983	2	T00385	672	25	65.8	200	2	E97360	NAD(P)-flavin oxi
600	26	68.4	1997	2	T30874	673	25	65.8	200	2	AF2578	hypothetical prote
601	26	68.4	2052	2	T37711	674	25	65.8	205	2	A71567	probable GMP kinas
602	26	68.4	2165	1	RNNZAZ	675	25	65.8	205	2	B84384	uroporphyrin-TII C
603	26	68.4	2175	1	GNNYBE	676	25	65.8	207	2	H84275	fuculose-1-phospha
604	26	68.4	2205	2	T08615	677	25	65.8	211	1	S74011	probable riboflavi
605	26	68.4	2569	2	T14164	678	25	65.8	213	1	AC0252	4-hydroxy-2-oxogl
606	26	68.4	2570	2	A35548	679	25	65.8	213	2	E72020	inorganic pyrophos
607	26	68.4	2870	2	H96974	680	25	65.8	215	2	D86605	60S ribosomal prot
608	26	68.4	3149	1	QOBE8	681	25	65.8	216	2	F90135	
609	25	65.8	56	2	E82317	682	25	65.8				
610	25	65.8	56	2	E89770	683	25	65.8				
611	25	65.8	70	2	D82598	684	25	65.8				
612	25	65.8	72	2	T49510	685	25	65.8				
613	25	65.8	73	2	T36840	686	25	65.8				

687	25	65.8	221	2	G87508	hydrolyase, haloacti	760	25	65.8	294	2	D87505	geranyltranstransf
688	25	65.8	221	2	C95009	potassium uptake p	761	25	65.8	294	2	T01459	hypothetical prote
689	25	65.8	223	2	B90178	translacton initia	762	25	65.8	295	2	AE1566	probable UDP-gluc
690	25	65.8	223	2	T30147	hypothetical prote	763	25	65.8	295	2	AE1566	kinase, GMP fami
691	25	65.8	225	1	ADPSCP	2-dehydro-3-deoxy-	764	25	65.8	296	2	E81299	hypothetical prote
692	25	65.8	225	2	F91294	2-component transc	765	25	65.8	296	2	A70973	hypothetical prote
693	25	65.8	225	2	A86136	bgly protein - Esc	766	25	65.8	297	1	S49348	cytochrome-c oxida
694	25	65.8	225	2	S56593	hypothetical prote	767	25	65.8	298	2	G98159	probable transcrip
695	25	65.8	225	2	D84049	hypothetical prote	768	25	65.8	298	2	AB3128	transcription regu
696	25	65.8	228	2	G97880	hypothetical prote	769	25	65.8	299	2	AB1871	recombination asso
697	25	65.8	238	2	G83700	probable transfe	770	25	65.8	299	2	F81149	LysR-family regula
698	25	65.8	233	2	T06915	secretory protein	771	25	65.8	300	2	A97607	hypothetical prote
699	25	65.8	233	2	G81307	two-component resp	772	25	65.8	303	2	AH0082	hypothetical prote
700	25	65.8	234	2	JH0483	two-component resp	773	25	65.8	303	2	S56552	probable transcrip
701	25	65.8	236	2	AE1387	ribonuclease H11 -	774	25	65.8	303	2	T34112	hypothetical prote
702	25	65.8	236	2	AG1762	probable Tail-like	775	25	65.8	305	2	B84779	hypothetical prote
703	25	65.8	238	2	B72320	ubiquinol-cytochri	776	25	65.8	307	2	T38206	cytochrome-c oxida
704	25	65.8	240	2	G72295	probable transcrip	777	25	65.8	308	2	F87283	probable transcrip
705	25	65.8	240	2	A84543	transcription acti	778	25	65.8	309	2	S57102	transcription regu
706	25	65.8	242	2	B86131	type III secretion	779	25	65.8	310	2	AI2848	Rdc protein NMA10
707	25	65.8	243	2	B97423	hypothetical prote	780	25	65.8	312	2	S51085	recombination asso
708	25	65.8	243	2	JQ0021	hypothetical prote	781	25	65.8	313	2	T29195	LysR-family regula
709	25	65.8	243	2	AB2640	cytochrome-c oxida	782	25	65.8	313	2	S46690	hypothetical prote
710	25	65.8	244	2	G91093	cytochrome-c oxida	783	25	65.8	315	1	KTECR	hypothetical prote
711	25	65.8	244	2	H71873	cytochrome-c oxida	784	25	65.8	315	1	H00842	probable phosphopr
712	25	65.8	245	2	DB1977	cytochrome-c oxida	785	25	65.8	315	2	E85700	methionyl-tRNA for
713	25	65.8	247	1	OBVQM	cytochrome-c oxida	786	25	65.8	315	2	AE0245	hypothetical prote
714	25	65.8	247	1	OBVQM	cytochrome-c oxida	787	25	65.8	315	2	AE0245	CDK-activating kin
715	25	65.8	247	1	OBVQM	cytochrome-c oxida	788	25	65.8	315	2	AE0720	hypothetical prote
716	25	65.8	247	2	S42736	cytochrome-c oxida	789	25	65.8	315	2	B84403	hypothetical prote
717	25	65.8	247	2	S42736	cytochrome-c oxida	790	25	65.8	318	2	G97625	ribose-phosphate d
718	25	65.8	247	2	S42736	cytochrome-c oxida	791	25	65.8	319	2	S71176	dipeptide ABC tran
719	25	65.8	249	2	G46637	20S proteasome sub	792	25	65.8	320	2	B69643	prsa (AE006059) (1
720	25	65.8	250	2	A82919	integrase-recombin	793	25	65.8	322	2	C83075	DNA-directed RNA p
721	25	65.8	251	1	OBVY2	cytochrome-c oxida	794	25	65.8	322	2	B83075	hypothetical prote
722	25	65.8	251	2	S45322	cytochrome-c oxida	795	25	65.8	323	2	B51109	octaprenyl-diphosp
723	25	65.8	251	2	A87651	conserved hypotet	796	25	65.8	323	2	B91137	octaprenyl diphosp
724	25	65.8	255	2	T25733	hypothetical prote	797	25	65.8	323	2	E85982	octaprenyl diphosp
725	25	65.8	255	2	T25733	conserved hypotet	798	25	65.8	323	2	AB0427	probable prsa prot
726	25	65.8	255	2	AC2829	streptomycin 3 - E	799	25	65.8	323	2	AH0903	ribose-phosphate d
727	25	65.8	258	2	AC2829	coat protein - gar	800	25	65.8	326	2	D70622	probable ribose-ph
728	25	65.8	259	2	S27916	hypothetical prote	801	25	65.8	327	2	H86939	oligopeptide trans
729	25	65.8	259	2	JQ2179	conserved hypotet	802	25	65.8	327	2	B69656	hypothetical prote
730	25	65.8	259	2	G82601	hypothetical prote	803	25	65.8	327	2	F72673	hypothetical prote
731	25	65.8	261	2	B64375	conserved hypotet	804	25	65.8	327	2	G84378	hypothetical prote
732	25	65.8	262	2	B85959	transcription regu	805	25	65.8	328	2	A85728	hypothetical prote
733	25	65.8	262	2	H81720	hypothetical prote	806	25	65.8	328	2	H90889	threonine ammonia-
734	25	65.8	264	2	AD3201	hypothetical prote	807	25	65.8	328	2	AB0897	ribose-phosphate p
735	25	65.8	266	2	S67182	hypothetical prote	808	25	65.8	330	2	JN0886	ribose-phosphate d
736	25	65.8	267	2	G70576	probable cyto prot	809	25	65.8	331	2	I39927	ABC transporter (b
737	25	65.8	269	2	B23957	streptomycin/spect	810	25	65.8	332	2	S76660	ribose-phosphate d
738	25	65.8	271	2	E97907	phosphomethylpyrim	811	25	65.8	332	2	E84374	ABC transporter (b
739	25	65.8	271	2	E97907	hypothetical prote	812	25	65.8	333	2	I39927	ribose-phosphate d
740	25	65.8	275	2	D95259	ABC transporter, A	813	25	65.8	334	2	S76660	ribose-phosphate d
741	25	65.8	275	2	G98124	hypothetical prote	814	25	65.8	335	2	E70453	malate dehydrogena
742	25	65.8	277	2	S75973	hypothetical prote	815	25	65.8	336	2	J50052	37K protein - Myo
743	25	65.8	280	2	A82185	glycerol-3-phospha	816	25	65.8	336	2	T21565	hypothetical prote
744	25	65.8	280	2	S47815	hypothetical prote	817	25	65.8	336	2	G84025	polysugar degradin
745	25	65.8	281	2	F86034	hypothetical prote	818	25	65.8	337	2	T52589	ribose-phosphate d
746	25	65.8	281	2	EBBET2	30-2K zinc-binding	819	25	65.8	339	2	T42081	probable transpos
747	25	65.8	283	2	E86766	transcription regu	820	25	65.8	340	2	C14711	conserved hypotet
748	25	65.8	285	1	S13870	ubiquinol-cytochro	821	25	65.8	341	2	T69149	C 4.2.3.9 aristolo
749	25	65.8	286	2	G75489	hypothetical prote	822	25	65.8	342	2	A45462	hypothetical prote
750	25	65.8	286	2	T02119	probable DnaJ prot	823	25	65.8	344	2	T23028	cell division cycl
751	25	65.8	286	2	E90053	DM-binding protei	824	25	65.8	346	2	S34652	geranyltranstransf
752	25	65.8	289	2	T44462	hypothetical prote	825	25	65.8	346	2	A56231	hypothetical prote
753	25	65.8	290	2	A97777	hypothetical prote	826	25	65.8	347	2	T11664	octaprenyl-diphosp
754	25	65.8	291	2	T49260	conserved hypotet	827	25	65.8	347	2	A70139	hypothetical prote
755	25	65.8	292	1	A40650	conserved hypotet	828	25	65.8	347	2	S36980	octaprenyl-diphosp
756	25	65.8	292	1	C82162	UTP-glucose-1-phos	829	25	65.8	348	2	S34494	hypothetical prote
757	25	65.8	293	2	EB4114	UTP-glucose-1-phos	830	25	65.8	348	2	AE0049	csa protein - Eug
758	25	65.8	293	2	C71703	hypothetical prote	831	25	65.8	350	2	AG1721	M48 peptidase fami
759	25	65.8	293	2	C71703	hypothetical prote	832	25	65.8	350	2	AG1721	uroporphyrinogen I



833	25	65.8	352	2	A75098	906	25	65.8	421	2	D86240	hypothetical prote
834	25	65.8	353	2	AD1351	907	25	65.8	421	2	T43406	cutlin-3 - fission
835	25	65.8	353	2	B89908	908	25	65.8	422	2	S22898	vitellogenin 2 - M
836	25	65.8	354	2	H95217	909	25	65.8	424	2	C82516	type I restriction
837	25	65.8	354	2	H98081	910	25	65.8	427	1	YKEC	citrate (si)-synh
838	25	65.8	354	2	T14980	911	25	65.8	427	2	A99722	citrate synthase [
839	25	65.8	355	2	G72325	912	25	65.8	427	2	H85572	citrate synthase [
840	25	65.8	355	2	S28523	913	25	65.8	427	2	AF0590	citrate synthase [
841	25	65.8	355	2	JN0746	914	25	65.8	427	2	AC0823	peptidase B (leucy
842	25	65.8	356	2	F71151	915	25	65.8	432	2	H82358	conserved hypothet
843	25	65.8	360	2	T51870	916	25	65.8	432	2	G69977	two-component sens
844	25	65.8	361	2	T27504	917	25	65.8	432	2	AI0351	peptidase B (lmpor
845	25	65.8	363	2	D64640	918	25	65.8	432	2	S63447	hypothetical prote
846	25	65.8	363	2	B83057	919	25	65.8	439	2	E97079	sugar-binding peri
847	25	65.8	365	2	S12091	920	25	65.8	441	2	H83763	6-phospho-beta-glu
848	25	65.8	366	1	B69640	921	25	65.8	441	2	T11256	retinoic acid rece
849	25	65.8	366	2	T31553	922	25	65.8	448	2	C98206	hypothetical trans
850	25	65.8	366	2	D97002	923	25	65.8	448	2	AF3080	transcription regu
851	25	65.8	367	1	UC1349	924	25	65.8	448	2	S57686	hypothetical prote
852	25	65.8	367	1	151426	925	25	65.8	450	2	S15675	globulin-2 precurs
853	25	65.8	367	2	C96537	926	25	65.8	451	2	E70013	hypothetical prote
854	25	65.8	367	2	E69072	927	25	65.8	451	2	A88641	hypothetical prote
855	25	65.8	367	2	AC1371	928	25	65.8	451	2	T30732	A-type inclusion b
856	25	65.8	367	2	A11740	929	25	65.8	455	2	T34155	hypothetical prote
857	25	65.8	367	2	T39313	930	25	65.8	456	2	E91052	probable peptidase
858	25	65.8	368	1	S42582	931	25	65.8	456	2	A85897	hypothetical prote
859	25	65.8	368	2	C90011	932	25	65.8	456	2	B65029	CAP59 protein - Cr
860	25	65.8	369	2	AB1537	933	25	65.8	458	2	A56055	protein C (activat
861	25	65.8	371	2	D83650	934	25	65.8	461	1	KXHU	hypothetical prote
862	25	65.8	371	2	S46408	935	25	65.8	461	2	B97228	hypothetical prote
863	25	65.8	372	2	T39244	936	25	65.8	463	2	T46290	mg2+ transporter m
864	25	65.8	375	2	H84666	937	25	65.8	465	2	AI1361	glutamine transpor
865	25	65.8	375	2	G70966	938	25	65.8	465	2	S73400	probable DEAD/DEAF
866	25	65.8	375	2	T52300	939	25	65.8	465	2	E96737	hypothetical prote
867	25	65.8	376	1	JE0023	940	25	65.8	467	2	T10025	hypothetical prote
868	25	65.8	376	2	AF0783	941	25	65.8	469	2	T35670	hypothetical prote
869	25	65.8	379	2	G82233	942	25	65.8	469	2	F70393	Xaa-His dipeptidas
870	25	65.8	379	2	H89327	943	25	65.8	470	2	AD1277	Xaa-His dipeptidas
871	25	65.8	382	2	A99992	944	25	65.8	470	2	AD1640	hypothetical prote
872	25	65.8	383	2	AG3312	945	25	65.8	470	2	S54089	hypothetical prote
873	25	65.8	386	2	T02640	946	25	65.8	476	2	S60927	glu-tRNAGln amidot
874	25	65.8	389	2	B44972	947	25	65.8	478	2	C97266	UDP-N-acetylmuram
875	25	65.8	389	2	D64337	948	25	65.8	481	2	AI1323	flagellar hook-ass
876	25	65.8	392	2	T36967	949	25	65.8	484	2	AF3329	cardiolipin synthe
877	25	65.8	392	2	S31228	950	25	65.8	486	2	A56145	cardiolipin synthe
878	25	65.8	394	2	D97411	951	25	65.8	486	2	AG0266	cardiolipin syntha
879	25	65.8	394	2	AF2629	952	25	65.8	486	2	D85705	cardiolipin synthe
880	25	65.8	395	2	F81343	953	25	65.8	486	2	E90847	cardiolipin synthe
881	25	65.8	395	2	E83367	954	25	65.8	486	2	AF0651	conserved hypothet
882	25	65.8	396	2	F70970	955	25	65.8	489	2	D98312	hypothetical prote
883	25	65.8	398	2	B70308	956	25	65.8	489	2	AG1384	hypothetical prote
884	25	65.8	398	2	H72660	957	25	65.8	491	2	AI1759	conserved hypothet
885	25	65.8	398	2	B90406	958	25	65.8	491	2	C72213	hypothetical prote
886	25	65.8	401	2	AC3454	959	25	65.8	497	2	AB4171	aromatic-L-amino-a
887	25	65.8	401	2	AB3082	960	25	65.8	500	1	DCDAAP	probable phosphoe
888	25	65.8	402	2	JCS151	961	25	65.8	506	2	S76981	cellulagin I sylv
889	25	65.8	402	2	S31196	962	25	65.8	511	2	S58399	envelope glycoprot
890	25	65.8	404	1	JCS150	963	25	65.8	512	1	VG1WTH	hypothetical prote
891	25	65.8	404	1	D82355	964	25	65.8	515	2	H83406	conserved hypothet
892	25	65.8	405	1	A64314	965	25	65.8	517	2	C69550	heat shock transcr
893	25	65.8	405	1	B36807	966	25	65.8	517	2	B40583	hypothetical prote
894	25	65.8	411	2	A65184	967	25	65.8	518	2	T05196	unspecific monooxy
895	25	65.8	411	2	B91220	968	25	65.8	519	2	S69989	RNA-directed RNA p
896	25	65.8	411	2	D86066	969	25	65.8	521	2	A29249	oligonucleotide ABC
897	25	65.8	411	2	T19195	970	25	65.8	521	2	S38539	characterized pro
898	25	65.8	412	2	H70915	971	25	65.8	524	2	B84028	hypothetical prote
899	25	65.8	412	2	F98204	972	25	65.8	525	2	C97170	heat shock transcr
900	25	65.8	414	2	G90173	973	25	65.8	527	2	T23740	heat shock transcr
901	25	65.8	414	2	T35625	974	25	65.8	529	2	A41138	sugar transport pr
902	25	65.8	416	2	A69023	975	25	65.8	536	2	T22856	hypothetical prote
903	25	65.8	416	2	T00813	976	25	65.8	547	2	A96667	
904	25	65.8	418	2	AF3316	977	25	65.8	552	2		
905	25	65.8	421	2		978	25	65.8		2		

```

979      25      65.8      555      2      A83327      probable AMP-dindi
980      25      65.8      555      2      C70177      beta-glucosidase h
981      25      65.8      559      2      B96448      hypothetical prote
982      25      65.8      561      2      T19604      hypothetical prote
983      25      65.8      564      2      C42523      A5SR protein - vac
984      25      65.8      564      2      J01792      Salp17R protein -
985      25      65.8      564      2      S52682      hypothetical prote
986      25      65.8      573      2      T48584      auxin-regulated pr
987      25      65.8      576      2      B69312      probable acyl-CoA
988      25      65.8      584      2      S76424      hypothetical prote
989      25      65.8      589      2      T44486      glutamine-fructose
990      25      65.8      593      1      K1ZPMN      protein kinase cdr
991      25      65.8      594      2      T48087      transport inhibito
992      25      65.8      599      2      S67495      huntingtin-associa
993      25      65.8      608      2      H72292      hypothetical prote
994      25      65.8      616      1      RRVOL1      probable RNA-direc
995      25      65.8      616      1      RRVOL1      probable RNA-direc
996      25      65.8      621      2      E64546      glucose inhibited
997      25      65.8      622      2      F81375      RNA polymerase sig
998      25      65.8      622      2      T49952      hypothetical prote
999      25      65.8      628      2      A81999      glucose inhibited
1000     25      65.8      629      2      C82598      glucose inhibited

```

## ALIGNMENTS

```

RESULT 1
D59433
C: elegans protein Z37093 homolog [imported] - hnamn
C:Species: Homo sapiens (man)
C:Date: 03-Jun-2002 #sequence_revision 03-Jun-2002 #text_change 03-Jun-2002
C:Accession: D59433; E59433
R:Nagase, T.; Seki, N.; Ishikawa, K.; Ohira, M.; Kawarabayashi, Y.; Ohara, O.; Tanaka, A.
DNA Res. 3, 321-329, 1996
A:Title: Prediction of the coding sequences of unidentified human genes. VI. The coding
A:Reference number: D59433; M0ID:97191544; PMID:9039502
A:Accession: D59433
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1165 <NAC>
A:Cross-references: GB:BA13212; PID:q1504026; PIDN:BA13212.1
R:Ohara, O.; Nagase, T.; Kikuno, R.; Nomura, N.
Submitted to Genbank, August 1996
A:Reference number: E59433
A:Accession: E59433
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1165 <STO2>
A:Cross-references: GB:BA13212; PID:q1504026; PIDN:BA13212.1

Query Match      94.7%; Score 36; DB 2; Length 1165;
Best Local Similarity 88.9%; Pred. No. 31;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY      1 VLXDDLLEA 9
      |||||||
Db      166 VLKDDLLEA 174

RESULT 2
C69226
type I restriction modification enzyme, subunit M - Methanobacterium thermoautotrophicum
C:Species: Methanobacterium thermoautotrophicum
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 29-Sep-1999
C:Accession: C69226
R:Smith, D.R.; Doucelte-Stamm, L.A.; Delouhery, C.; Lee, H.; Dubois, J.; Aldredge, T.;
; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.
; K.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: funct
A:Reference number: A69000; M0ID:9803514; PMID:9371463

```

```

A:Accession: C69226
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-616 <MTH>
A:Cross-references: GB:A600086; GB:A600066; M0ID:92622025; PIDN:AAB85440.1; PID:9262
C:Genetics:
A:Gene: MTH942
A:Start codon: GTG
C:Superfamily: type I site-specific deoxyribonuclease chain hsdM

Query Match      89.5%; Score 34; DB 2; Length 616;
Best Local Similarity 77.8%; Pred. No. 40;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY      1 VLXDDLLEA 9
      |||||||
Db      465 VLKDDLLEA 473

```

```

RESULT 3
C89075
protein K04A8.9 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
C:Accession: C89075
R:Anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating bio
A:Reference number: A75000; M0ID:99069613; PMID:9851916
A:Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/projects/C-
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999;
A:Accession: C89075
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-174 <STO>
A:Cross-references: GB:chr_V; PIDN:AC48055.1; PID:q1658358; GSPDB:GN00023; CESP:K04A
C:Genetics:
A:Gene: K04A8.9
A:Map position: 5

```

```

Query Match      86.8%; Score 33; DB 2; Length 174;
Best Local Similarity 77.8%; Pred. No. 16;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY      1 VLXDDLLEA 9
      |||||||
Db      129 VLKDDLLEA 137

```

```

RESULT 4
G86538
leucyl aminopeptidase A [imported] - Chlamydomonas reinhardtii (strain J138)
C:Species: Chlamydomonas reinhardtii
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
C:Accession: G86538
R:Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.;
Nucleic Acids Res. 28, 2311-2314, 2000
A:Title: Comparison of whole genome sequences of Chlamydomonas reinhardtii J138.
A:Reference number: A86491; M0ID:20330349; PMID:10871362
A:Accession: G86538
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-499 <STO>
A:Cross-references: GB:BA000008; M0ID:98978757; PIDN:BA898593.1; GSPDB:GN00142
C:Experimental source: strain J138
C:Genetics:
A:Gene: pepA
C:Superfamily: cytosol aminopeptidase

Query Match      86.8%; Score 33; DB 2; Length 499;
Best Local Similarity 77.8%; Pred. No. 51;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 VLXDDLEA 9  
||:|||||  
Db 395 VLAEEDLEA 403

## RESULT 5

leucine aminopeptidase CP0370 [imported] - Chlamydomophila pneumoniae (strains CML029 and G72083)  
M:Alternate names: leucyl aminopeptidase A  
C:Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae  
C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 11-May-2000  
C:Accession: G72083; F81583  
R:Kallan, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.; Nature Genet. 21, 385-389, 1999  
A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.  
A:Reference number: A72000; MUID:99206606; PMID:10192388  
A:Accession: G72083  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-499 <ARN>  
A:Cross-references: GB:AE001623; GB:AE001363; NID:g4376662; PIDN:AD18529.1; PID:g437666  
A:Experimental source: strain CML029  
R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey, C.; Dodson, R.; Swinn, W.; Nelson, W.; Deboy, R.; Kolonay, J.; McClarty, G.; Salzberg, Nucleic Acids Res. 28, 1397-1406, 2000  
A:Title: Genome sequences of Chlamydia trachomatis Mopn and Chlamydia pneumoniae AR39.  
A:Reference number: AB1500; MUID:20150255; PMID:10684935  
A:Accession: F81583  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-499 <REA>  
A:Cross-references: GB:AE002199; GB:AE002161; NID:g7189293; PIDN:AAE38219.1; PID:g718929  
A:Experimental source: strain AR39, HL cells  
C:Genetics:  
A:Gene: pepA; CP0370  
C:Superfamily: cytosol aminopeptidase

Query Match 86.8%; Score 33; DB 2; Length 499;  
Best Local Similarity 77.8%; Pred. No. 51;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLXDDLEA 9  
||:|||||  
Db 395 VLAEEDLEA 403

## RESULT 6

nitrogen regulatory protein P-II - Rhodobacter sphaeroides  
C:Species: Rhodobacter sphaeroides  
C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 26-Aug-1999  
C:Accession: S33180  
R:Zinchenko, V.V.; Churin, U.N.; Shestopalov, V.I.; Shestakov, S.V.  
A:Description: The EMBL data library, April 1993  
A:Reference number: S33180  
A:Accession: S33180  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-112 <ZIN>  
A:Cross-references: EMBL:X71659; NID:g297112; PIDN:CA50650.1; PID:g809751  
A:Experimental source: strain 2R  
C:Genetics:  
A:Gene: glnB  
A:Start codon: GTG  
C:Function:  
A:Description: important for the control of glutamine synthase; in nitrogen-limiting con-  
dition, these events are reversed  
GS Note: P-II indirectly controls the transcription of the GS gene (glnA) by preventing N  
C:Superfamily: regulatory protein P-II

C:Keywords: phosphoprotein; signal transduction  
E:51/Binding site: UMP (Tyr) (covalent) #status predicted

Query Match 84.2%; Score 32; DB 2; Length 112;  
Best Local Similarity 66.7%; Pred. No. 16;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLXDDLEA 9  
||:|||||  
Db 64 VLADDMVEA 72

## RESULT 7

hypothetical protein SPAC1486.03c [imported] - fission yeast (Schizosaccharomyces pombe T50072)  
C:Species: Schizosaccharomyces pombe  
C:Date: 09-Jun-2000 #sequence\_revision 09-Jun-2000 #text\_change 09-Jun-2000  
C:Accession: T50072  
R:McDougal, R.C.; Rajandream, M.A.; Barrell, B.G.; Seeger, K.; Harris, D.  
submitted to the EMBL data library, December 1999  
A:Reference number: 225034  
A:Accession: T50072  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-797 <MCD>  
A:Cross-references: EMBL:AL133357; PIDN:CA62413.1; GSPDB:GN00066; SPDB:SPAC1486.03c  
A:Experimental source: strain 972h(-); cosmid c1486  
C:Genetics:  
A:Gene: SPDB:SPAC1486.03c  
A:Map position: 1

Query Match 84.2%; Score 32; DB 2; Length 797;  
Best Local Similarity 77.8%; Pred. No. 144+02;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 VLXDDLEA 9  
||:|||||  
Db 81 VTADDDLEA 89

## RESULT 8

sugar ABC transporter, permease protein homolog lin1843 [imported] - Listeria innocua AB1663  
C:Species: Listeria innocua  
C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 14-Dec-2001  
C:Accession: AB1663  
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloec, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurgey, O.; Entian, K.D.; Fsihl, D.; Jones, L.M.; Karst, U.  
Science 294, 849-852, 2001  
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkai, G.; Madueno, E.; Mollournam, A.; Ok, C.; Schluteler, T.; Simoes, N.; Tlerriz, A.; Vazquez-Boland, J.A.; Voss, H.; Mehla, A:Title: Comparative genomics of Listeria species.  
A:Reference number: AB1077; MUID:21537279; PMID:11679669  
A:Accession: AB1663  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-276 <GLA>  
A:Cross-references: GB:AL592022; PIDN:CA697074.1; PID:g16414345; GSPDB:GN00178  
A:Experimental source: strain C11p11262  
C:Genetics:  
A:Gene: lin1843  
C:Superfamily: maltose transport protein malC

Query Match 81.6%; Score 31; DB 2; Length 276;  
Best Local Similarity 75.0%; Pred. No. 68;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LXDDELEA 9  
||:|||||  
Db 165 ISDDLEA 172

RESULT 9  
AD1291  
sugar ABC transporter, permease protein homolog lmo1732 [imported] - *Listeria monocytogenes*  
C:Species: *Listeria monocytogenes*  
C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 14-Dec-2001  
C:Accession: AD1291  
R:Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bioecker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussauguet, O.; Enlitan, K.D.; Fsihl, H. Science 294, 849-852, 2001  
A:Authors: Kreft, J., Kuhn, M.; Kunst, F.; Kurapkai, G.; Madueno, E.; Maitournam, A.; Mok, C.; Schluteter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.; Title: Comparative genomics of *Listeria* species.  
A:Reference number: AB1077; MUID:21537279; PMID:11679669  
A:Accession: AD1291  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-276 <GLA>  
A:Cross-references: GB:NC\_003210; PIDN:CAC99810.1; PID:g16411186; GSPDB:GN00177  
A:Experimental source: strain EGD-e  
C:Genetics:  
A:Gene: lmo1732  
C:Superfamily: maltose transport protein malG

Query Match 81.6%; Score 31; DB 2; Length 276;  
Best Local Similarity 75.0%; Pred. No. 68;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 2 VLXDDLEA 9 :|||||  
DB 165 ILADLDLEA 172

RESULT 10  
DB1930  
probable UTP-glucose-1-phosphate uridylyltransferase NMA0848 [imported] - *Neisseria meningitidis*  
C:Species: *Neisseria meningitidis*  
C:Date: 05-May-2000 #sequence\_revision 05-May-2000 #text\_change 02-Feb-2001  
C:Accession: DB1930  
R:Parhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churche, C.; Klee, S.R.; Morel, H.; Holroyd, S.; Jagsels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream, A.; Title: Complete DNA sequence of a serogroup A strain of *Neisseria meningitidis* Z2491.  
A:Reference number: AB1775; MUID:20222556; PMID:10761919  
A:Accession: DB1930  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-289 <PAR>  
A:Cross-references: GB:AL162754; GB:AL157959; NID:g7379424; PIDN:CAB84129.1; PID:g737956  
A:Experimental source: serogroup A, strain Z2491  
C:Genetics:  
A:Gene: galU; NMA0848  
C:Superfamily: *Escherichia coli* UTP-glucose-1-phosphate uridylyltransferase

Query Match 81.6%; Score 31; DB 2; Length 289;  
Best Local Similarity 55.6%; Pred. No. 72;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
QY 1 VLXDDLEA 9 :|||||  
DB 130 ILADLDLEA 138

RESULT 11  
C81175  
UTP-glucose-1-phosphate uridylyltransferase NMB0638 [imported] - *Neisseria meningitidis*  
C:Species: *Neisseria meningitidis*  
C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 19-Jan-2001  
C:Accession: C81175  
R:Reitlin, H.; Saunders, N.J.; Heiselberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.; Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.; Ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignani, V.; Pizsa, M. Science 287, 1809-1815, 2000

A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; A:Title: Complete genome sequence of *Neisseria meningitidis* serogroup B strain MC58.  
A:Reference number: AB1000; MUID:20175755; PMID:10710307  
A:Accession: C81175  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-289 <TEV>  
A:Cross-references: GB:AE002419; GB:AE002098; NID:g7225863; PIDN:AAF41061.1; PID:g722  
A:Experimental source: serogroup B, strain MC58  
C:Genetics:  
A:Gene: NMB0638  
C:Superfamily: *Escherichia coli* UTP-glucose-1-phosphate uridylyltransferase

Query Match 81.6%; Score 31; DB 2; Length 289;  
Best Local Similarity 55.6%; Pred. No. 72;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
QY 1 VLXDDLEA 9 :|||||  
DB 130 ILADLDLEA 138

RESULT 12  
A28180  
carbamoyl-phosphate synthase (glutamine-hydrolyzing) (EC 6.3.5.5) large chain - *Methanobacterium thermoautotrophicum*  
C:Species: *Methanobacterium thermoautotrophicum*  
C:Date: 20-Jun-1989 #sequence\_revision 20-Jun-1989 #text\_change 07-Aug-1998  
C:Accession: A28180  
R:Morris, C.J.; Reeve, J.N.  
J. Bacteriol. 170, 3125-3130, 1988  
A:Title: Conservation of structure in the human gene encoding argininosuccinate synthetase.  
A:Reference number: A91873; MUID:88257029; PMID:3133361  
A:Accession: A28180  
A:Molecule type: DNA  
A:Residues: 1-397 <MOR>  
C:Genetics:  
A:Gene: canB  
C:Superfamily: biotin carboxylase homology  
C:Keywords: arginine biosynthesis; ligase; pyrimidine nucleotide biosynthesis  
F:1-308/Domain: biotin carboxylase homology <BCH>

Query Match 81.6%; Score 31; DB 2; Length 397;  
Best Local Similarity 66.7%; Pred. No. 1e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
QY 1 VLXDDLEA 9 :|||||  
DB 67 ILIDDFLEA 75

RESULT 13  
G87793  
protein C27A12.7 [imported] - *Caenorhabditis elegans*  
C:Species: *Caenorhabditis elegans*  
C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 10-May-2001  
C:Accession: G87793  
R:anonymous, The C. elegans Sequencing Consortium.  
Science 287, 2012-2018, 1998  
A:Title: Genome sequence of the nematode *C. elegans*: a platform for investigating bio  
A:Reference number: A75000; MUID:99069613; PMID:9851916  
A:Note: see websites genome.wustl.edu/gsc/C\_elegans/ and www.sanger.ac.uk/projects/C\_elegans/  
A:Accession: G87793  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-497 <STO>  
A:Cross-references: GB:chr\_1; PIDN:AA93644.1; PID:g2105479; GSPDB:GN00019; CESP:C27A  
C:Genetics:  
A:Gene: C27A12.7  
A:Map position: 1

Query Match 81.6%; Score 31; DB 2; Length 497;

Best Local Similarity 77.8%; Pred. No. 1.3e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 VLXDDLEA 9  
|||:|||||

Db 55 VLNDLLEA 63

RESULT 14

B83794 hypothetical protein BH1154 [imported] - Bacillus halodurans (strain C-125)

C:Species: Bacillus halodurans

C:Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 15-Jun-2001

C:Accession: B83794

R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hirai

Nucleic Acids Res. 28, 4317-4331, 2000

A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and

A:Reference number: A83650; MUID:20512582; PMID:11058132

A:Accession: B83794

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-750 <STO>

A:Cross-references: GB:AP001511; GB:BA000004; NID:g10173727; PIDN:BA04873.1; GSPDB:GN00

A:Experimental source: strain C-125

C:Genetics:

A:Gene: BH1154

Query Match 81.6%; Score 31; DB 2; Length 750;  
Best Local Similarity 66.7%; Pred. No. 2.1e+02;

Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 VLXDDLEA 9  
|||:|||||

Db 581 ILTDDLEA 589

RESULT 15

T13647

hypothetical protein 95B7.8 - fruit fly (Drosophila melanogaster)

C:Species: Drosophila melanogaster

C:Date: 13-Aug-1999 #sequence\_revision 13-Aug-1999 #text\_change 17-Nov-2000

C:Accession: T13647

R:Ferraz, C.; Vidal, S.; Brun, C.; Bucheton, A.; Demaille, J.G.

submitted to the EMBL Data Library, April 1999

A:Description: Sequencing the distal X chromosome of Drosophila melanogaster.

A:Reference number: Z17694

A:Accession: T13647

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1042 <FER>

A:Cross-references: EMBL:AL021728; PIDN:CA16820.1

C:Genetics:

A:Cross-references: FLYBase:FBgn0000376

A:Introns: 94/3; 833/3; 937/3

A:Note: EG:95B7.8

C:Superfamily: fruit fly hypothetical protein EG\_95B7.8

Query Match 81.6%; Score 31; DB 2; Length 1042;  
Best Local Similarity 77.8%; Pred. No. 3e+02;

Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 VLXDDLEA 9  
|||:|||||

Db 595 VLFPDDLEA 603

RESULT 16

T26975

hypothetical protein Y47H9C.12 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999

C:Accession: T26975

R:Harris, B.

submitted to the EMBL Data Library, October 1998

A:Reference number: Z20293

A:Accession: T26975

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-72 <WIL>

A:Cross-references: EMBL:AL032657; PIDN:CAA21742.1; GSPDB:GN00019; CESP:Y47H9C.12

A:Experimental source: clone Y47H9C

C:Genetics:

A:Gene: CESP:Y47H9C.12

A:Map position: 1

A:Introns: 21/2

Query Match 78.9%; Score 30; DB 2; Length 72;  
Best Local Similarity 75.0%; Pred. No. 25;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 VLXDDLE 8  
|||:|||||

Db 61 VLGDDEME 68

RESULT 17

AB3246 transposase tnp [imported] - Agrobacterium tumefaciens (strain C58, Dupont) plasmid T

C:Species: Agrobacterium tumefaciens

C:Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 11-Jan-2002

C:Accession: AB3246

R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Moo

erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayvin, T.; Levy, R.; Li, M.; McCl

; Karp, P.; Romero, P.; Zhang, S.

Science 294, 2317-2323, 2001

A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kam

ster, E.W.

A:Title: The genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A:Reference number: AB2577; PMID:11743193

A:Accession: AB3246

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-302 <KUP>

A:Cross-references: GB:AF008690; PIDN:AL46384.1; PID:g17744176; GSPDB:GN00189

A:Experimental source: strain C58 (Dupont)

C:Genetics:

A:Gene: tnp

A:Genome: plasmid

Query Match 78.9%; Score 30; DB 2; Length 302;  
Best Local Similarity 75.0%; Pred. No. 1.2e+02;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 LKXDDLEA 9  
|||:|||||

Db 59 LTDDVLEA 66

RESULT 18

T48014

serine/threonine protein kinase-like protein - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 02-Sep-2000

C:Accession: T48014

R:Rieger, M.; Mueller-Auer, S.; Zipp, M.; Schaefer, M.; Mewes, H.W.; Lemcke, K.; Maye

submitted to the Protein Sequence Database, February 2000

A:Reference number: Z24482

A:Accession: T48014

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-361 <RIE>

A:Cross-references: EMBL:AL138651

A:Experimental source: cultivar Columbia; BAC clone T17J13

C:Genetics:

A:Map position: 3

A:Introns: 29/4; 112/3; 155/1; 246/2; 290/3; 319/3  
 A:Note: T17J13.180  
 C:Superfamily: kinase-related transforming protein; protein kinase homology

Query Match 78.9%; Score 30; DB 2; Length 361;  
 Best Local Similarity 55.6%; Pred. No. 1.5e+02;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLXDDLEA 9  
 : | | | | |  
 Db 57 ILADELIFA 65

## RESULT 19

E75221

Type I restriction modification enzyme, chain m. PAB2149 - *Pyrococcus abyssi* (strain Ors  
 C:Species: *Pyrococcus abyssi*  
 C:Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 20-Jun-2000

C:Accession: E75221  
 R:anonymous, Genoscope

A:Description: The EMBL Data Library, July 1999

A:Reference number: A75001

A:Accession: E75221

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-623 <KMW>

A:Cross-references: GB:AJ248283; GB:AL096836; NID:95457433; PIDN:CAB49220.1; PID:9545772

A:Experimental source: strain Orsay

C:Genetics:

A:Gene: PAB2149

C:Superfamily: type I site-specific deoxyribonuclease chain hsdM

Query Match 78.9%; Score 30; DB 2; Length 623;  
 Best Local Similarity 55.6%; Pred. No. 2.7e+02;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLXDDLEA 9  
 : | | | | |  
 Db 472 IVEDDLIFA 480

## RESULT 20

D82674

Tom-dependent receptor for iron transport Xf1496 [imported] - *Xylella fastidiosa* (strain  
 C:Species: *Xylella fastidiosa*  
 C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 20-Aug-2000

C:Accession: D82674

R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen

A:Title: The genome sequence of the plant pathogen *Xylella fastidiosa*.

A:Reference number: A82515; MUID:20365717; PMID:10910347

A:Note: for a complete list of authors see reference number A59328 below

A:Accession: D82674

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-663 <SIM>

A:Cross-references: GB:AE003979; GB:AE003849; NID:99106520; PIDN:AAF84305.1; GSPDB:GN001

A:Experimental source: strain 945c

R:Simpon, A.J.G.; Relnach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A

Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carrato, D.M.; Carter, H

submitted to GenBank, June 2000

A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Fromt

J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig

chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, B

, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Miyaki, C.Y.;

Rodrigues, V.; Rosa, A.J.; de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak

A:Autours: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir

M.; Tsuhako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z

A:Reference number: A59328

A:Contents: annotation

C:Genetics:  
 A:Gene: Xf1496

Query Match 78.9%; Score 30; DB 2; Length 683;  
 Best Local Similarity 75.0%; Pred. No. 3e+02;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLXDDLE 8  
 : | | | | |  
 Db 65 VLSDDLIQ 72

## RESULT 21

I38728

epidermal growth factor receptor kinase substrate - human  
 C:Species: *Homo sapiens* (man)

C:Date: 29-May-1998 #sequence\_revision 29-May-1998 #text\_change 09-Jun-2000

C:Accession: I38728

R:Mong, W.T.; Carlomagno, F.; Druck, T.; Barletta, C.; Croce, C.M.; Huebner, K.; Krau

OncoGene 9, 3057-3061, 1994

A:Title: Evolutionary conservation of the Eps8 gene and its mapping to human chromoso

A:Reference number: I38728; MUID:94366758; PMID:8084614

A:Accession: I38728

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-822 <RES>

A:Cross-references: EMBL:U12535; NID:9530822; PID:9530823

C:Genetics:

A:Gene: Eps8

C:Superfamily: SH3 homology

Query Match 78.9%; Score 30; DB 2; Length 822;  
 Best Local Similarity 75.0%; Pred. No. 3.7e+02;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLXDDLE 8  
 : | | | | |  
 Db 552 VLKDDILE 559

## RESULT 22

J70669

helicase II-like protein, B962L - African swine fever virus  
 C:Species: African swine fever virus, ASFV

C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 19-Jan-2001

C:Accession: J70669

R:Itanetz, R.J.; Rodriguez, J.M.; Bournnell, M.; Rodriguez, J.F.; Vinuela, E.

Gene 134, 161-174, 1993

A:Title: Two putative African swine fever virus helicases similar to yeast 'DEAH' pre

A:Reference number: J70665; MUID:94085774; PMID:8262374

A:Accession: J70669

A:Molecule type: DNA

A:Residues: 1-962 <YAN>

A:Cross-references: GB:U18466; NID:9780375; PIDN:AAA65302.1; PID:9780442

C:Superfamily: African swine fever virus probable helicase II B962L

C:Keywords: Amp; nucleotide binding; P-loop

F:165-63/Region: nucleotide-binding motif A (P-loop)

F:165-168/Region: nucleotide-binding motif B

F:167-170/Region: DEAH motif

Query Match 78.9%; Score 30; DB 1; Length 962;  
 Best Local Similarity 75.0%; Pred. No. 4.4e+02;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLXDDLE 8  
 : | | | | |  
 Db 584 LLADLLE 591

RESULT 23

AB2018

hypothetical protein al11696 [imported] - *Nostoc* sp. (strain PCC 7120)

C:Species: *Nostoc* sp.



Query Match 76.3%; Score 29; DB 2; Length 101;  
 Best Local Similarity 85.7%; Pred. No. 59;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 2 LXDILLE 8  
 | | | | |  
 Db 69 LDDILLE 75

## RESULT 28

735859

hypothetical protein SC9B1.14c - Streptomyces coelicolor

C:Species: Streptomyces coelicolor

C&gt;Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 21-Jan-2000

C:Accession: J35859

R:Saunders, D.C.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.

submitted to the EMBL Data Library, April 1999

A:Reference number: 221591

A:Accession: J35859

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-119 &lt;SAUV&gt;

A:Cross-references: EMBL:AI049727; PIDN:CA041560.1; GSPDB:GN00070; SC0EDB:SC9B1.14c

A:Experimental source: strain A3(2)

C:Genetics:

A:Gene: SC0EDB:SC9B1.14c

C:Superfamily: Mycobacterium tuberculosis hypothetical protein RV3363c

Query Match 76.3%; Score 29; DB 2; Length 119;  
 Best Local Similarity 66.7%; Pred. No. 70;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 VLXDILLEA 9  
 | | | | |  
 Db 80 LILSDILLEA 88

## RESULT 29

03BCEA

ydbd protein - Escherichia coli (strain K-12)

C:Species: Escherichia coli

C&gt;Date: 31-Dec-1989 #sequence\_revision 31-Dec-1989 #text\_change 01-Mar-2002

C:Accession: B91904; C91903; C64793; C32046; D32047

R:Li, J.; Duncan, K.; Malsb, C.T.

J. Bacteriol. 171, 791-796, 1989

A:Title: Nucleotide sequence of a cluster of Escherichia coli enterobactin biosynthesis

A:Reference number: A91904; MUID:89123155; PMID:2521622

A:Accession: B91904

A:Molecule type: DNA

A:Residues: 1-137 &lt;LTD&gt;

A:Cross-references: GB:M24148; NID:9304949; PIDN:AAA16104.1; PID:9450383

J. Bacteriol. 171, 784-790, 1989

A:Title: Nucleotide sequence and transcriptional organization of the Escherichia coli en

A:Reference number: A91903; MUID:89123154; PMID:2521621

A:Accession: C91903

A:Molecule type: DNA

A:Residues: 1-137 &lt;NAH&gt;

A:Cross-references: GB:M24143; NID:9341118; PIDN:AAA76837.1; PID:9522184

R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; C

Science 277, 1453-1462, 1997

A:Title: The complete genome sequence of Escherichia coli K-12.

A:Reference number: A64720; MUID:97426617; PMID:9278503

A:Accession: C64793

A:Molecule type: DNA

A:Status: nucleic acid sequence not shown; translation not shown

A:Residues: 1-137 &lt;BLAT&gt;

A:Cross-references: GB:AE000165; GB:U00096; NID:91786808; PIDN:AACT3698.1; PID:91786813;

A:Experimental source: strain K-12, substrain MG1655

C:Genetics:

A:Gene: ydbd

A:Map position: 13 min

C:Superfamily: Escherichia coli ydbd protein

Query Match 76.3%; Score 29; DB 1; Length 137;  
 Best Local Similarity 75.0%; Pred. No. 82;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 2 LXDILLEA 9  
 | | | | |  
 Db 31 LGDDVLEA 38

## RESULT 30

D90708

hypothetical protein EC50636 [imported] - Escherichia coli (strain O157:H7, substrain

C:Species: Escherichia coli

C&gt;Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 03-Aug-2001

C:Accession: D90708

R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C

gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.

DNA Res. 8, 11-22, 2001

A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and g

A:Reference number: A99629; MUID:21156231; PMID:11258796

A:Accession: D90708

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-137 &lt;HAY&gt;

A:Cross-references: GB:BA000007; PIDN:BA034059.1; PID:913360094; GSPDB:GN00154

A:Experimental source: strain O157:H7, substrain RIMD 0509552

C:Genetics:

A:Gene: EC50636

C:Superfamily: Escherichia coli ydbd protein

Query Match 76.3%; Score 29; DB 2; Length 137;  
 Best Local Similarity 75.0%; Pred. No. 82;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 2 LXDILLEA 9  
 | | | | |  
 Db 31 LGDDVLEA 38

## RESULT 31

H85558

hypothetical protein ydbd [imported] - Escherichia coli (strain O157:H7, substrain ED

C:Species: Escherichia coli

C&gt;Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-Sep-2001

C:Accession: H85558

R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May

Miller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Diallanta, E.; Potamoustis, K.; Apoda

Nature 409, 529-533, 2001

A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A:Reference number: A85480; MUID:21074935; PMID:11206551

A:Accession: H85558

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-137 &lt;STO&gt;

A:Cross-references: GB:AE005174; NID:912513490; PIDN:AA654932.1; GSPDB:GN00145; UWGP:

A:Experimental source: strain O157:H7, substrain EDL933

C:Genetics:

A:Gene: ydbd

C:Superfamily: Escherichia coli ydbd protein

Query Match 76.3%; Score 29; DB 2; Length 137;  
 Best Local Similarity 75.0%; Pred. No. 82;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 2 LXDILLEA 9  
 | | | | |  
 Db 31 LGDDVLEA 38

## RESULT 32

A10575



conserved hypothetical protein STY0643 [imported] - *Salmonella enterica* subsp. *enterica*  
 C:Species: *Salmonella enterica* subsp. *enterica* serovar Typh  
 A:Note: this species has also been called *Salmonella typhi*  
 C:Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 27-Nov-2001  
 C:Accession: A10575  
 R:Parfhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moule, S.; O'Gaora, P.  
 Nature 413, 848-852, 2001  
 A:Authors: Parry, C.; Quall, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Nature 413, 848-852, 2001  
 A:Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* serovar Typhimurium  
 A:Reference number: AB0502; PMID:11677608  
 A:Accession: A10575  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-137 <PAR>  
 A:Cross-references: GB:AL513382; PIDN:CAD05075.1; PID:G16501850; GSPDB:GN00176  
 C:Genetics:  
 A:Gene: STY0643  
 C:Superfamily: *Escherichia coli* ybdB protein

Query Match 76.3%; Score 29; DB 2; Length 137;  
 Best Local Similarity 75.0%; Pred. No. 82;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 LKDDLEA 9  
 Db 31 LGDDVLEA 38

## RESULT 33

R5HS30  
 ribosomal protein L30 [validated] - *Halococcus marismortui*  
 N:Alternate names: ribosomal protein HL16; ribosomal protein HL20  
 C:Species: *Halococcus marismortui*  
 C:Date: 31-Mar-1991 #sequence\_revision 31-Dec-1992 #text\_change 31-Mar-2000  
 C:Accession: S16543; S06845; B28926; T46808  
 R:Scholezen, T.; Arndt, E.  
 Mol. Gen. Genet. 228, 70-80, 1991  
 A:Title: Organization and nucleotide sequence of ten ribosomal protein genes from the re  
 A:Reference number: S16535; MOID:91360093; PMID:1832208  
 A:Accession: S16543  
 A:Molecule type: DNA  
 A:Residues: 1-154 <SCH>  
 A:Cross-references: EMBL:X58395; NID:G48860; PIDN:CA41292.1; PID:G48869  
 R:Hatakeyama, T.; Kaufmann, F.; Schroeter, B.; Hatakeyama, T.  
 Eur. J. Biochem. 185, 685-693, 1989  
 A:Title: Primary structures of five ribosomal proteins from the archaeobacterium *Halobact*  
 A:Reference number: S06844; MOID:90076190; PMID:2591382  
 A:Accession: S06845

A:Molecule type: protein  
 A:Residues: 1-82, L', 84-147, 149-154 <HAT>  
 A:Note: the source is designated as *Halobacterium marismortui*  
 A:Note: the protein is designated as ribosomal protein HL20  
 R:Walsh, M.J.; McDougall, J.; Wilmann-Jibbold, B.  
 Biochemistry 27, 6867-6876, 1988  
 A:Title: Extended N-terminal sequencing of proteins of archaeobacterial ribosomes plotted  
 A:Reference number: A28926; MOID:89062418; PMID:3196689  
 A:Accession: B28926  
 A:Molecule type: protein  
 A:Residues: 1-30 <WAL>  
 A:Note: the protein is designated as ribosomal protein L16  
 C:Genetics:  
 A:Gene: Hmal30  
 C:Superfamily: rat ribosomal protein L7  
 C:Keywords: protein biosynthesis; ribosome  
 F:1-154/Product: ribosomal protein L30 #status experimental <MAT>

Query Match 76.3%; Score 29; DB 1; Length 154;  
 Best Local Similarity 100.0%; Pred. No. 94;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 DDLEA 9

Db 147 DDLEA 152

## RESULT 34

T20267  
 hypothetical protein C56A3.2 - *Caenorhabditis elegans*  
 C:Species: *Caenorhabditis elegans*  
 C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 21-Jan-2000  
 C:Accession: T20267  
 R:Sim, M.  
 submitted to the EMBL Data Library, July 1996  
 A:Reference number: Z19244  
 A:Accession: T20267  
 A:Status: preliminary  
 A:Molecule type: translated from GB/EMBL/DBJ  
 A:Residues: 1-156 <WIL>  
 A:Cross-references: EMBL:Z77655; PIDN:CA601136.1; GSPDB:GN00023; CESP:C56A3.2  
 A:Experimental source: clone C56A3  
 C:Genetics:  
 A:Gene: CESP:C56A3.2  
 A:Map position: 5  
 A:Insertions: 93/3  
 C:Superfamily: *Caenorhabditis* hypothetical protein C40H1.5

Query Match 76.3%; Score 29; DB 2; Length 156;  
 Best Local Similarity 100.0%; Pred. No. 95;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 DDLEA 9  
 Db 52 DDLEA 57

## RESULT 35

D72293  
 hypothetical protein - *Thermotoga maritima* (strain MSB8)  
 C:Species: *Thermotoga maritima*  
 C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 28-Jul-2000  
 C:Accession: D72293  
 R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwin, M.L.; Dodson, R.J.; Haft, D.H.; Hic  
 Garrelt, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, C.M.  
 Nature 399, 323-329, 1999  
 A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome  
 A:Reference number: A72200; MOID:99287316; PMID:10360571  
 A:Accession: D72293  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-178 <ARN>  
 A:Cross-references: GB:AE001770; GB:AE000512; NID:G4981658; PIDN:AAD36203.1; PID:G498  
 A:Experimental source: strain MSB8  
 C:Genetics:  
 A:Gene: TM1127  
 C:Superfamily: *Thermotoga maritima* hypothetical protein TM1127

Query Match 76.3%; Score 29; DB 2; Length 178;  
 Best Local Similarity 75.0%; Pred. No. 11e+02;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 LKDDLEA 9  
 Db 58 LADDDLEA 65

## RESULT 36

A12271  
 transcription regulator all3728 [imported] - *Nostoc* sp. (strain PCC 7120)  
 C:Species: *Nostoc* sp.  
 A:Note: *Nostoc* sp. strain PCC 7120 is a synonym of *Anabaena* sp. strain PCC 7120  
 C:Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 30-Jun-2002  
 C:Accession: A12271  
 R:Kaneko, T.; Nakamura, Y.; Molk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Irigu

Nakazaki, N.; Shimpou, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.  
 DNA Res. 8, 205-213, 2001  
 A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium *Ana*  
 A:Reference number: AB1807; MUID:21595285; PMID:11759840  
 A:Accession: AI2271  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-192 <KUR>  
 A:Cross-references: GB:BA000019; PIDN:BAF75427.1; PID:g17132862; GSPDB:GN00179  
 A:Experimental source: strain PCC 7120  
 C:Genetics:  
 A:Gene: all3728

Query Match 76.3%; Score 29; DB 2; Length 192;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 DDLLEA 9  
 |||||  
 Db 61 DDLLEA 66

RESULT 37  
 G83538  
 Probable oxidoreductase PA0853 [imported] - *Pseudomonas aeruginosa* (strain PA01)  
 C:Species: *Pseudomonas aeruginosa*  
 C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
 C:Accession: G83538  
 R:Stover, C.K.; Pham, X.O.; Ewlin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Br  
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardi, K.; Lim,  
 Loay, S.; Olson, M.V.  
 Nature 406, 959-964, 2000  
 A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic patho  
 A:Reference number: AB2950; MUID:20437337; PMID:10964043  
 A:Accession: G83538  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-207 <STO>  
 A:Cross-references: GB:AE004520; GB:AE004091; NID:g9946745; PIDN:AG04242.1; GSPDB:GN001  
 A:Experimental source: strain PA01  
 C:Genetics:  
 A:Gene: PA0853

Query Match 76.3%; Score 29; DB 2; Length 207;  
 Best Local Similarity 77.8%; Pred. No. 1.3e+02;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 VLXDDLEA 9  
 |||||  
 Db 73 VLXDDLEA 81

RESULT 38  
 B65140  
 Hypothetical 26.3 kD protein in gntR-ggt intergenic region - *Escherichia coli* (strain K-  
 C:Species: *Escherichia coli*  
 C:Date: 12-Sep-1997 #sequence\_revision 17-Sep-1997 #text\_change 01-Mar-2002  
 C:Accession: B65140  
 R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; CC  
 A.; Rose, D.J.; Mau, B.; Shao, Y.  
 Science 277, 1453-1462, 1997  
 A:Title: The complete genome sequence of *Escherichia coli* K-12.  
 A:Reference number: A64720; MUID:97426617; PMID:9278503  
 A:Accession: B65140  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-231 <BLAT>  
 A:Cross-references: GB:AE000420; GB:U00096; NID:g1789840; PIDN:AACT6464.1; PID:g1789847;  
 A:Experimental source: strain K-12, substrain MG1655  
 C:Genetics:  
 A:Gene: yhhw  
 C:Superfamily: conserved hypothetical protein sl11773

Query Match 76.3%; Score 29; DB 2; Length 231;  
 Best Local Similarity 55.6%; Pred. No. 1.5e+02;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 VLXDDLEA 9  
 |||||  
 Db 42 VINDVIEA 50

RESULT 39  
 H91164  
 Hypothetical protein ECs4288 [imported] - *Escherichia coli* (strain O157:H7, substrain  
 C:Species: *Escherichia coli*  
 C:Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 03-Aug-2001  
 C:Accession: H91164  
 R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C  
 gasawara, N.; Yasunaga, T.; Kihara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.  
 DNA Res. 8, 11-22, 2001  
 A:Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and 9  
 A:Reference number: A99629; MUID:21156231; PMID:11258796  
 A:Accession: H91164  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-231 <HAY>  
 A:Cross-references: GB:BA000007; PIDN:BA837711.1; PID:g13363762; GSPDB:GN00154  
 A:Experimental source: strain O157:H7, substrain RIMD 0509952  
 C:Genetics:  
 A:Gene: ECs4288  
 C:Superfamily: conserved hypothetical protein sl11773

Query Match 76.3%; Score 29; DB 2; Length 231;  
 Best Local Similarity 55.6%; Pred. No. 1.5e+02;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 VLXDDLEA 9  
 |||||  
 Db 42 VINDVIEA 50

RESULT 40  
 H86010  
 Hypothetical protein yhhw [imported] - *Escherichia coli* (strain O157:H7, substrain ED  
 C:Species: *Escherichia coli*  
 C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-Sep-2001  
 C:Accession: H86010  
 R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May  
 iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apoda  
 Nature 409, 529-533, 2001  
 A:Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.  
 A:Reference number: A85480; MUID:21074935; PMID:11206551  
 A:Accession: H86010  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-231 <STO>  
 A:Cross-references: GB:AE005174; NID:g12518089; PIDN:AA658548.1; GSPDB:GN00145; UMGCP:  
 A:Experimental source: strain O157:H7, substrain EDL933  
 C:Genetics:  
 A:Gene: yhhw  
 C:Superfamily: conserved hypothetical protein sl11773

Query Match 76.3%; Score 29; DB 2; Length 231;  
 Best Local Similarity 55.6%; Pred. No. 1.5e+02;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 VLXDDLEA 9  
 |||||  
 Db 42 VINDVIEA 50

RESULT 41  
 D86725  
 Hypothetical protein yibB [imported] - *Lactococcus lactis* subsp. *lactis* (strain IL140  
 C:Species: *Lactococcus lactis* subsp. *lactis*

C>Date: 23-Mar-2001 #sequence\_revision 23-Mar-2001 #text\_change 03-Aug-2001  
 C:Accession: D66725  
 R:Polinton, A.; Winkler, P.; Mauger, S.; Jallion, O.; Malarme, K.; Weissbach, J.; Ehrlich  
 Genome Res. 11, 731-753, 2001  
 A>Title: The complete genome sequence of the lactic acid bacterium *Lactococcus lactis* s8  
 A:Reference number: A66625; MUID:21235186; PMID:11337471  
 A:Accession: D66725  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-244 <STO>  
 A:Cross-references: GB:AE005176; PID:g12723725; PIDN:AAK04902.1; GSPDB:GM00146  
 A:Experimental source: strain IL1403  
 C:Genetics:  
 A:Gene: y1bB

Query Match 76.3%; Score 29; DB 2; Length 244;  
 Best Local Similarity 100.0%; Pred. No. 1.0e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 DDLLEA 9  
 |||||  
 DB 71 DDLLEA 76

## RESULT 42

Cytochrome-c oxidase (EC 1.9.3.1) chain II - yeast (*Candida glabrata*) mitochondrion  
 C:Species: mitochondrion *Candida glabrata*  
 C>Date: 07-May-1995 #sequence\_revision 14-Jul-1995 #text\_change 07-Dec-1999  
 C:Accession: S45438  
 R:Clark-Walker, G.D.; Weller, G.F.  
 J. Mol. Evol. 38, 593-601, 1994

A>Title: The structure of the small mitochondrial DNA of *Kluyveromyces thermotolerans* is  
 A:Reference number: S45322; MUID:94365843; PMID:8083884  
 A:Accession: S45438  
 A:Status: translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-251 <CLA>  
 A:Cross-references: EMBL:X69430; NID:9509746; PIDN:CAA49205.1; PID:g728536  
 C:Genetics:  
 A:Gene: COXII  
 A:Genome: mitochondrion  
 A:Genetic code: SGC2  
 C:Superfamily: cytochrome-c oxidase chain II; cytochrome-c oxidase chain II homology  
 C:Keywords: copper; electron transfer; heme; membrane-associated complex; mitochondrial  
 ein

F:25-239/Domain: cytochrome-c oxidase chain II homology <CO2>  
 F:186-221-225-239/Binding site: copper 1 (His, Cys, Cys, Met) #status predicted  
 F:221-223/225-239/Binding site: copper 2 (Cys, Glu, Cys, His) #status predicted  
 F:223/Binding site: magnesium (Glu) (shared with chain I) #status predicted

Query Match 76.3%; Score 29; DB 2; Length 251;  
 Best Local Similarity 75.0%; Pred. No. 1.6e+02;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLXDDLE 8  
 |||||  
 DB 147 VLXDDLE 154

## RESULT 43

D69158  
 sensory transduction regulatory protein - *Methanobacterium thermoautotrophicum* (strain D  
 C:Species: *Methanobacterium thermoautotrophicum*  
 C>Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 22-Oct-1999  
 C:Accession: D69158  
 R:Smith, D.R.; Doucette-Stamm, L.A.; Delonghery, C.; Lee, H.; Dubois, J.; Aldredge, T.;  
 ; Olm, D.; Spadafora, R.; Vitale, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.  
 K. S. Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.  
 J. Bacteriol. 179, 7135-7155, 1997

A>Title: Complete genome sequence of *Methanobacterium thermoautotrophicum* Delta H: funct  
 A:Reference number: A69000; MUID:98037514; PMID:9371463  
 A:Accession: D69158

A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-277 <MRH>  
 A:Cross-references: GB:AE000829; GB:AE000666; NID:92621512; PIDN:AA84953.1; PID:9262  
 A:Experimental source: strain Delta H  
 C:Genetics:  
 A:Gene: MTH447  
 C:Superfamily: response regulator homology  
 C:Keywords: phosphoprotein  
 F:4-114/Domain: response regulator homology <RRH>  
 F:53/Binding site: phosphate (Asp) (covalent) #status predicted

Query Match 76.3%; Score 29; DB 2; Length 277;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 DDLLEA 9  
 |||||  
 DB 187 DDLLEA 192

## RESULT 44

S72323  
 ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome c1 precursor - yeast (*Kluyv*  
 C:Species: *Kluyveromyces fragilis* var. *lactis* *Candida sphaerica*  
 C>Date: 04-Feb-1998 #sequence\_revision 13-Feb-1998 #text\_change 03-Jun-2002  
 C:Accession: S72323  
 R:Delella, Y.; Horvathova, K.; van der Aart, O.J.M.; Zonneveld, B.J.M.; Steensma, H.Y  
 Curr. Genet. 30, 145-150, 1996

A>Title: Isolation and molecular analysis of the gene for cytochrome c1 from *Kluyve*  
 A:Reference number: S72323; MUID:96304292; PMID:8660461  
 A:Accession: S72323  
 A:Status: nucleic acid sequence not shown  
 A:Molecule type: DNA  
 A:Residues: 1-292 <GBE>  
 A:Cross-references: EMBL:X95899; NID:91209427; PIDN:CAA65144.1; PID:g1209428  
 A:Experimental source: strain JBD100  
 A:Note: the source is designated as *Kluyveromyces lactis*  
 C:Genetics:  
 A:Gene: CYT1  
 A:Map position: 6  
 A:Genome: nuclear  
 A:Superfamily: cytochrome c1 heme protein; cytochrome c1 heme protein homology  
 C:Keywords: chromoprotein; electron transfer; heme; iron; membrane-associated complex

F:1-46/Domain: transit peptide (mitochondrion) #status predicted <TP>  
 F:47-292/Product: cytochrome c1 #status predicted <MAT>  
 F:55-281/Domain: cytochrome c1 heme protein homology <CIH>  
 F:254-271/Domain: transmembrane #status predicted <TM>  
 F:86-89/Binding site: heme (Cys) (covalent) #status predicted  
 F:90-210/Binding site: heme iron (His, Met) (axial ligands) #status predicted

Query Match 76.3%; Score 29; DB 2; Length 292;  
 Best Local Similarity 75.0%; Pred. No. 1.9e+02;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLXDDLE 8  
 |||||  
 DB 213 VLXDDLE 220

## RESULT 45

S44847  
 K06H7.7 protein - *Caenorhabditis elegans*  
 C:Species: *Caenorhabditis elegans*  
 C>Date: 14-Sep-1994 #sequence\_revision 12-May-1995 #text\_change 23-Mar-2001  
 C:Accession: S44847  
 R:Revello, A.D.  
 Submitted to the EMBL Data Library, May 1993

A:Description: Sequence of the *C. elegans* cosmid K06H7.  
 A:Reference number: S44620  
 A:Accession: S44847  
 A:Status: preliminary

A:Molecule type: DNA  
 A:Residues: 1-322 <FAV>  
 A:Cross-references: EMBL:L15314; NID:g289690; PID:g289697  
 C:Genetics:  
 A:Introns: 56/1; 116/1; 142/2; 181/3; 281/3

Query Match 76.3%; Score 29; DB 2; Length 322;  
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 2 LXDLEA 8  
 Db 221 LADLEA 227

RESULT 46  
 E90949  
 heat shock protein MsbA [imported] - Escherichia coli (strain O157:H7, substrain RMD 05  
 C:Species: Escherichia coli  
 C:Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 18-Jul-2001  
 C:Accession: E90949  
 R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.  
 Gaswara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.  
 DNA Res. 8, 11-22, 2001  
 A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genc  
 A:Reference number: A99629; MUID:21156231; PMID:11258796  
 A:Accession: E90949  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-323 <HAY>  
 A:Cross-references: GB:BA000007; PIDN:BA935988.1; PID:g13362033; GSPDB:GN00154  
 A:Experimental source: strain O157:H7, substrain RMD 0509952  
 C:Genetics:  
 A:Gene: ECS2565

Query Match 76.3%; Score 29; DB 2; Length 323;  
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 4 DDLLEA 9  
 Db 267 DDLLEA 272

RESULT 47  
 A85798  
 suppressor of htrB, heat shock protein [imported] - Escherichia coli (strain O157:H7, su  
 C:Species: Escherichia coli  
 C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-Sep-2001  
 C:Accession: A85798  
 R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew  
 Miller, L.; Grothbeck, E.J.; Davis, N.W.; Jim, A.; Dimatanta, E.; Potamousis, K.; Apodaca,  
 Nature 409, 529-533, 2001  
 A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
 A:Reference number: A85480; MUID:21074935; PMID:11206551  
 A:Accession: A85798  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-323 <STO>  
 A:Cross-references: GB:AE005174; NID:g12515911; PIDN:AA656845.1; GSPDB:GN00145; UWGP:228  
 A:Experimental source: strain O157:H7, substrain EDL933  
 C:Genetics:  
 A:Gene: msbA

Query Match 76.3%; Score 29; DB 2; Length 323;  
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 4 DDLLEA 9  
 Db 267 DDLLEA 272

RESULT 48  
 A42608  
 (Kdo)2-(lauroyl)-lipid IVa acyltransferase (EC 2.3.1.-) - Escherichia coli (strain K-  
 N:Alternate names: membrane-bound lytic transglycosylase  
 C:Species: Escherichia coli  
 C:Date: 04-Mar-1993 #sequence\_revision 18-Nov-1994 #text\_change 01-Mar-2002  
 C:Accession: A42608; 154979; 664947  
 R:Karow, M.; Georgopoulos, C.  
 J. Bacteriol. 174, 702-710, 1992  
 A:Title: Isolation and characterization of the Escherichia coli msbA gene, a multico  
 A:Reference number: A42608; MUID:92121107; PMID:1732206  
 A:Accession: A42608  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-323 <KAR>  
 A:Cross-references: GB:M87660; NID:g146876; PIDN:AAA96706.1; PID:g146877  
 A:Experimental source: strain K-12, substrain W3110  
 A:Note: Sequence extracted from NCBI backbone (NCBI:77416, NCBI:77418)  
 R:Engel, H.; Smink, A.J.; van Wijngaarden, L.; Keck, W.  
 J. Bacteriol. 174, 6394-6403, 1992  
 A:Title: Murein-metabolizing enzymes from Escherichia coli: existence of a second lyt  
 A:Reference number: I54979; MUID:93015688; PMID:1356966  
 A:Accession: I54979  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-323 <RES>  
 A:Cross-references: GB:M87660; NID:g146876; PIDN:AAA96706.1; PID:g146877  
 R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;  
 .A.; Rose, D.J.; Mau, B.; Shao, Y.  
 Science 277, 1453-1462, 1997  
 A:Title: The complete genome sequence of Escherichia coli K-12.  
 A:Reference number: A64720; MUID:97426617; PMID:9278503  
 A:Accession: G64947  
 A:Status: nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-323 <BLAT>  
 A:Cross-references: GB:AE000279; GB:U00096; NID:g1788154; PIDN:AACT4925.1; PID:g17881  
 A:Experimental source: strain K-12, substrain MG1655  
 C:Genetics:  
 A:Gene: msbA; mlt

Query Match 76.3%; Score 29; DB 2; Length 323;  
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 4 DDLLEA 9  
 Db 267 DDLLEA 272

RESULT 49  
 AC2367  
 glucose-1-phosphate thymidyltransferase [imported] - Nostoc sp. (strain PCC 7120)  
 C:Species: Nostoc sp.  
 A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
 C:Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 30-Jun-2002  
 C:Accession: AC2367  
 R:Kaneko, T.; Nakamura, Y.; Maki, C.P.; Kurita, T.; Sasamoto, S.; Matsumoto, A.; Iriju  
 Nakazaki, N.; Shimizu, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata  
 DNA Res. 8, 205-213, 2001  
 A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium  
 A:Reference number: AB1807; MUID:21595285; PMID:11759840  
 A:Accession: AC2367  
 A:Status: preliminary  
 A:Molecule type: DNA

A:Residues: 1-358 <KUR>  
 A:Cross-references: GB:BA000019; PIDN:BA876190.1; PID:g17133627; GSPDB:GN00179  
 A:Experimental source: strain PCC 7120  
 C:Genetics:  
 A:Gene: alr4491  
 C:Superfamily: mannose-1-phosphate guanylyltransferase

Query Match 76.3%; Score 29; DB 2; Length 358;  
 Best Local Similarity 100.0%; Pred. No. 2.4e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 DDLLEA 9  
 |||||  
 Db 227 DDLLEA 232

## RESULT 50

B69852  
 Probable NADH2 dehydrogenase (EC 1.6.99.3) yj1D - Bacillus subtilis  
 C:Species: Bacillus subtilis  
 C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 03-Jun-2002  
 C:Accession: B69852  
 R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertet  
 C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chc  
 A.; Ehrlich, S.D.; Emmerison, P.T.; Ertlan, K.D.; Errington, J.; Fabret, C.; Ferrari, E.  
 Nature 390, 249-256, 1997  
 A:Authors: Foulger, D.; Filtz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gall  
 tech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.  
 A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue  
 A:Authors: Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portet  
 Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon  
 A:Authors: Schleich, S.; Schroeter, R.; Scofield, F.; Sekiguchi, J.; Sekowska, A.; Ser  
 akuchl, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,  
 T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K  
 A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.  
 A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.  
 A:Reference number: A69580; M01D:98044033; PMID:9384377  
 A:Accession: B69852  
 A:Status: nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-392 <KUN>  
 A:Cross-references: GB:299110; GB:AL009126; NID:g2633472; PIDN:CAB13086.1; PID:g2633583  
 A:Experimental source: strain 168  
 C:Genetics:  
 A:Gene: yj1D  
 C:Superfamily: NADH dehydrogenase  
 C:Keywords: electron transfer; FAD; flavoprotein; NAD; oxidoreductase

Query Match 76.3%; Score 29; DB 1; Length 392;  
 Best Local Similarity 75.0%; Pred. No. 2.6e+02;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 VLXDDLE 8  
 |||||  
 Db 205 VLXDDLE 212

Search completed: December 27, 2002, 14:42:30  
 Job time : 46 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 27, 2002, 14:40:34 ; Search time 11 Seconds  
(without alignments)  
33.935 Million cell updates/sec

Title: US-09-489-760A-1  
Sequence: 1 VLXDDLLEA 9

Scoring table: BLOSUM62  
Gap 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 1000 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	33	86.8	499	1	AMPB_CHLPP
2	32	84.2	112	1	GLNB_RHOSH
3	31	81.6	398	1	CARB_METMA
4	30	78.9	1073	1	CARB_METMA
5	29	76.3	822	1	EPSS_HUMAN
6	29	76.3	137	1	YEDB_ECOLI
7	29	76.3	154	1	RL30_HALMA
8	29	76.3	231	1	YHNM_ECO57
9	29	76.3	231	1	YHNM_ECOLI
10	29	76.3	251	1	COX2_CANGA
11	29	76.3	292	1	CYL_KLUFA
12	29	76.3	322	1	YMYT_CAEEL
13	29	76.3	323	1	MSBB_ECOLI
14	29	76.3	391	1	YJLD_BACSU
15	29	76.3	392	1	VATH_DROME
16	29	76.3	440	1	Y284_AQUAE
17	29	76.3	495	1	YSO2_CAEEL
18	29	76.3	496	1	AMWB_TRIRP
19	29	76.3	606	1	LEPA_BORBU
20	29	76.3	1441	1	CPSA_MOUSE
21	29	76.3	1442	1	CPSA_HUMAN
22	29	76.3	1444	1	FEER_BACST
23	29	76.3	81	1	FEER_BACST
24	29	76.3	81	1	FEER_BACST
25	29	76.3	124	1	RL7_RALSO
26	29	76.3	133	1	NIKR_ECOLI
27	29	76.3	133	1	NIKR_SALTY
28	29	76.3	249	1	PKSI_BACSU
29	29	76.3	252	1	VTM1_IABAN
30	29	76.3	252	1	VTM1_IACKB
31	29	76.3	252	1	VTM1_IAPOW
32	29	76.3	252	1	VTM1_IAPPR
33	29	76.3	252	1	VTM1_IAPPW

34	28	73.7	252	1	VTM1_IAMAN	P08381 influenza a
35	28	73.7	252	1	VTM1_IADDO	P03486 influenza a
36	28	73.7	252	1	VTM1_IAMIL	P05777 influenza a
37	28	73.7	252	1	VTM1_IATZL	P05776 influenza a
38	28	73.7	257	1	ECHE_MYCLE	P07137 mycobacteri
39	28	73.7	282	1	F3J4_YEAST	P32907 saccharomyc
40	28	73.7	282	1	Y32K_BNYVG	P19231 beet necrot
41	28	73.7	343	1	G3P_METUA	P05846 methanococ
42	28	73.7	434	1	PEPB_HAEIN	P58474 haemophilus
43	28	73.7	434	1	PEPB_PASMT	P06166 pasteurella
44	28	73.7	441	1	MAIR_FUSMR	P06901 fusobacteri
45	28	73.7	550	1	STR_CONGL	P35868 corynebacte
46	28	73.7	558	1	MNT2_YEAST	P33059 saccharomyc
47	28	73.7	735	1	DBH4_MOUSE	P31660 mus musculu
48	28	73.7	910	1	YK69_YEAST	P36165 saccharomyc
49	28	73.7	3519	1	OL56_STRAT	P00717 streptomyce
50	28	73.7	3829	1	SACS_HUMAN	P09214 homo sapien
51	28	73.7	3850	1	SACS_MOUSE	P09168 mus musculu
52	27	71.1	111	1	GLNB_BRAJA	P14179 bradyrhizob
53	27	71.1	119	1	RL7A_HALMA	P12743 haloarcula
54	27	71.1	139	1	RPE6_ASEFM	P08399 african swi
55	27	71.1	155	1	NRDG_HAEIN	P45080 haemophilus
56	27	71.1	156	1	ATPX_OPOSI	P00823 odontella s
57	27	71.1	179	1	APT_HELPY	P02199 helicobacte
58	27	71.1	179	1	APT_HELPY	P02199 helicobacte
59	27	71.1	239	1	SPSF_BACSU	P39626 bacillus su
60	27	71.1	261	1	YGFQ_ECOLI	P52045 escherichia
61	27	71.1	295	1	AMPW_PYRAB	P09474 pyrococcus
62	27	71.1	297	1	PUR7_MYCLE	P08361 mycobacteri
63	27	71.1	300	1	MNTA_BACHD	P07143 saccharomyc
64	27	71.1	309	1	CYL_YEAST	P07143 saccharomyc
65	27	71.1	325	1	ATP3_ARATH	P06250 arabidopsid
66	27	71.1	325	1	YCDW_ECOLI	P75913 escherichia
67	27	71.1	329	1	THD2_ECOLI	P05792 escherichia
68	27	71.1	331	1	DDH_ZYMO	P30799 zymomonas m
69	27	71.1	333	1	MRAN_ZYMO	P09699 zymomonas m
70	27	71.1	338	1	MTH2_HAEPA	P15446 haemophilus
71	27	71.1	401	1	CHUR_BACTN	P02550 bacteroides
72	27	71.1	415	1	ANM3_MOUSE	P09221 mus musculu
73	27	71.1	429	1	YDIS_ECOLI	P77337 escherichia
74	27	71.1	438	1	PUR2_PYRHO	P08061 pyrococcus
75	27	71.1	438	1	SYN_THERM	P54263 thermus the
76	27	71.1	465	1	RBL_DAPSP	P28397 daphniophyl
77	27	71.1	512	1	ANM3_HUMAN	P06678 homo sapien
78	27	71.1	513	1	VGA_BPPIX	P07928 bacterioph
79	27	71.1	522	1	VGA_BPPIX	P07928 bacterioph
80	27	71.1	528	1	ANM3_RAT	P07467 rattus norv
81	27	71.1	548	1	TRML_SCHPO	P09804 schizosacch
82	27	71.1	550	1	STR_MYCLE	P45840 mycobacteri
83	27	71.1	550	1	STR_MYCLE	P45840 mycobacteri
84	27	71.1	578	1	YC20_METUA	P058617 methanococ
85	27	71.1	603	1	LEPA_STYNS	P74751 synechocyst
86	27	71.1	631	1	CNG3_MOUSE	P09168 mus musculu
87	27	71.1	645	1	CNG3_CHICK	P090980 gallus gall
88	27	71.1	663	1	CNG2_BOVIN	P03041 bos taurus
89	27	71.1	664	1	CNG2_MOUSE	P062398 mus musculu
90	27	71.1	664	1	CNG2_MOUSE	P062398 mus musculu
91	27	71.1	664	1	CNG2_RABIT	P08718 oryctolagus
92	27	71.1	682	1	CNG2_RAT	P00195 rattus norv
93	27	71.1	683	1	CNG1_ICTPU	P35934 ictalurus p
94	27	71.1	684	1	CNG1_MOUSE	P062927 h cgmpr-gate
95	27	71.1	684	1	CNG1_MOUSE	P29974 h cgmpr-gate
96	27	71.1	690	1	CNG1_BOVIN	P29974 h cgmpr-gate
97	27	71.1	691	1	CNG1_BOVIN	P00194 b cgmpr-gate
98	27	71.1	694	1	CNG1_CANPA	P028279 c cgmpr-gate
99	27	71.1	704	1	CNG3_HUMAN	P016281 homo sapien
100	27	71.1	704	1	NEUL_HUMAN	P09474 homo sapien
101	27	71.1	704	1	NEUL_PIG	P02038 sus scrofa
102	27	71.1	706	1	NEUL_RABIT	P42675 oryctolagus
103	27	71.1	735	1	CNG3_BOVIN	P029441 bos taurus
104	27	71.1	735	1	CNG1_CHICK	P090805 gallus gall
105	27	71.1	821	1	EPSS_MOUSE	P00550 mus musculu
106	27	71.1	985	1	OPPF_MYCPN	P75551 mycoplasma
					DPOL_HSVIL	P28859 ictaluriid h

107	27	71.1	1011	1	SECA_PEA	041062	pisum sativ	180	26	66.4	465	1	RBL_PLAAC	P28441	platanus oc
108	27	71.1	1021	1	SECA_ARATH	095910	aradiopsis	181	26	66.4	465	1	RBL_TROAR	P28456	trochanodur oc
109	27	71.1	1036	1	SECA_SEPTOL	Q36795	spiniacia ol	182	26	66.4	468	1	GINA_MENCA	P15124	methylococc
110	27	71.1	1041	1	SYL_MYCTU	Q10765	mycobacteri	183	26	66.4	478	1	HLVD_PASNA	P15334	pasteurella
111	27	71.1	1059	1	SYL_MYCLE	Q9x7e5	mycobacteri	184	26	66.4	478	1	HLVD_PASNA	P15334	pasteurella
112	27	71.1	1070	1	CARB_METAC	08tuy4	methanosc	185	26	66.4	486	1	CUS_BUCAL	P57361	buchnera ap
113	27	71.1	1238	1	RPOB_MYCTU	P47766	mycobacteri	186	26	66.4	495	1	AMYB_SOYBN	P10538	glycine max
114	27	71.1	1238	1	YN13_YEAST	P53840	saccharomyc	187	26	66.4	498	1	AMYB_IPOHA	P10537	ipomoea bat
115	27	71.1	1597	1	SOL_DROME	P27398	drosophila	188	26	66.4	505	1	ALAT_SCHRO	P10703	schizosach
116	27	71.1	1616	1	RROO_TMYV	P03386	tobacco mos	189	26	66.4	505	1	ALAT_SCHRO	P10703	schizosach
117	27	71.1	1616	1	RROO_TMYV	Q9y908	tobacco mos	190	26	66.4	505	1	ALAT_SCHRO	P10703	schizosach
118	27	71.1	1616	1	RROO_TMYV	Q9y908	tobacco mos	191	26	66.4	505	1	ALAT_SCHRO	P10703	schizosach
119	27	71.1	1616	1	RROO_TMYV	P30738	tobacco mos	192	26	66.4	505	1	ALAT_SCHRO	P10703	schizosach
120	27	71.1	1616	1	RROO_TMYV	Q93058	tobacco mos	193	26	66.4	505	1	ALAT_SCHRO	P10703	schizosach
121	27	71.1	1616	1	RROO_TMYV	Q93058	tobacco mos	194	26	66.4	505	1	ALAT_SCHRO	P10703	schizosach
122	27	71.1	1616	1	RROO_TMYV	Q93058	tobacco mos	195	26	66.4	505	1	ALAT_SCHRO	P10703	schizosach
123	27	71.1	1616	1	RROO_TMYV	Q93058	tobacco mos	196	26	66.4	505	1	ALAT_SCHRO	P10703	schizosach
124	27	71.1	1616	1	RROO_TMYV	Q93058	tobacco mos	197	26	66.4	505	1	ALAT_SCHRO	P10703	schizosach
125	27	71.1	1616	1	RROO_TMYV	Q93058	tobacco mos	198	26	66.4	505	1	ALAT_SCHRO	P10703	schizosach
126	27	71.1	1616	1	RROO_TMYV	Q93058	tobacco mos	199	26	66.4	505	1	ALAT_SCHRO	P10703	schizosach
127	27	71.1	1616	1	RROO_TMYV	Q93058	tobacco mos	200	26	66.4	505	1	ALAT_SCHRO	P10703	schizosach
128	27	71.1	1616	1	RROO_TMYV	Q93058	tobacco mos	201	26	66.4	505	1	ALAT_SCHRO	P10703	schizosach
129	27	71.1	1616	1	RROO_TMYV	Q93058	tobacco mos	202	26	66.4	505	1	ALAT_SCHRO	P10703	schizosach
130	27	71.1	1616	1	RROO_TMYV	Q93058	tobacco mos	203	26	66.4	505	1	ALAT_SCHRO	P10703	schizosach
131	27	71.1	1616	1	RROO_TMYV	Q93058	tobacco mos	204	26	66.4	505	1	ALAT_SCHRO	P10703	schizosach
132	27	71.1	1616	1	RROO_TMYV	Q93058	tobacco mos	205	26	66.4	505	1	ALAT_SCHRO	P10703	schizosach
133	27	71.1	1616	1	RROO_TMYV	Q93058	tobacco mos	206	26	66.4	505	1	ALAT_SCHRO	P10703	schizosach
134	27	71.1	1616	1	RROO_TMYV	Q93058	tobacco mos	207	26	66.4	505	1	ALAT_SCHRO	P10703	schizosach
135	27	71.1	1616	1	RROO_TMYV	Q93058	tobacco mos	208	26	66.4	505	1	ALAT_SCHRO	P10703	schizosach
136	27	71.1	1616	1	RROO_TMYV	Q93058	tobacco mos	209	26	66.4	505	1	ALAT_SCHRO	P10703	schizosach
137	27	71.1	1616	1	RROO_TMYV	Q93058	tobacco mos	210	26	66.4	505	1	ALAT_SCHRO	P10703	schizosach
138	27	71.1	1616	1	RROO_TMYV	Q93058	tobacco mos	211	26	66.4	505	1	ALAT_SCHRO	P10703	schizosach
139	27	71.1	1616	1	RROO_TMYV	Q93058	tobacco mos	212	26	66.4	505	1	ALAT_SCHRO	P10703	schizosach
140	27	71.1	1616	1	RROO_TMYV	Q93058	tobacco mos	213	26	66.4	505	1	ALAT_SCHRO	P10703	schizosach
141	27	71.1	1616	1	RROO_TMYV	Q93058	tobacco mos	214	26	66.4	505	1	ALAT_SCHRO	P10703	schizosach
142	27	71.1	1616	1	RROO_TMYV	Q93058	tobacco mos	215	26	66.4	505	1	ALAT_SCHRO	P10703	schizosach
143	27	71.1	1616	1	RROO_TMYV	Q93058	tobacco mos	216	26	66.4	505	1	ALAT_SCHRO	P10703	schizosach
144	27	71.1	1616	1	RROO_TMYV	Q93058	tobacco mos	217	26	66.4	505	1	ALAT_SCHRO	P10703	schizosach
145	27	71.1	1616	1	RROO_TMYV	Q93058	tobacco mos	218	26	66.4	505	1	ALAT_SCHRO	P10703	schizosach
146	27	71.1	1616	1	RROO_TMYV	Q93058	tobacco mos	219	26	66.4	505	1	ALAT_SCHRO	P10703	schizosach
147	27	71.1	1616	1	RROO_TMYV	Q93058	tobacco mos	220	26	66.4	505	1	ALAT_SCHRO	P10703	schizosach
148	27	71.1	1616	1	RROO_TMYV	Q93058	tobacco mos	221	26	66.4	505	1	ALAT_SCHRO	P10703	schizosach
149	27	71.1	1616	1	RROO_TMYV	Q93058	tobacco mos	222	26	66.4	505	1	ALAT_SCHRO	P10703	schizosach
150	27	71.1	1616	1	RROO_TMYV	Q93058	tobacco mos	223	26	66.4	505	1	ALAT_SCHRO	P10703	schizosach
151	27	71.1	1616	1	RROO_TMYV	Q93058	tobacco mos	224	26	66.4	505	1	ALAT_SCHRO	P10703	schizosach
152	27	71.1	1616	1	RROO_TMYV	Q93058	tobacco mos	225	26	66.4	505	1	ALAT_SCHRO	P10703	schizosach
153	27	71.1	1616	1	RROO_TMYV	Q93058	tobacco mos	226	26	66.4	505	1	ALAT_SCHRO	P10703	schizosach
154	27	71.1	1616	1	RROO_TMYV	Q93058	tobacco mos	227	26	66.4	505	1	ALAT_SCHRO	P10703	schizosach
155	27	71.1	1616	1	RROO_TMYV	Q93058	tobacco mos	228	26	66.4	505	1	ALAT_SCHRO	P10703	schizosach
156	27	71.1	1616	1	RROO_TMYV	Q93058	tobacco mos	229	26	66.4	505	1	ALAT_SCHRO	P10703	schizosach
157	27	71.1	1616	1	RROO_TMYV	Q93058	tobacco mos	230	26	66.4	505	1	ALAT_SCHRO	P10703	schizosach
158	27	71.1	1616	1	RROO_TMYV	Q93058	tobacco mos	231	26	66.4	505	1	ALAT_SCHRO	P10703	schizosach
159	27	71.1	1616	1	RROO_TMYV	Q93058	tobacco mos	232	26	66.4	505	1	ALAT_SCHRO	P10703	schizosach
160	27	71.1	1616	1	RROO_TMYV	Q93058	tobacco mos	233	26	66.4	505	1	ALAT_SCHRO	P10703	schizosach
161	27	71.1	1616	1	RROO_TMYV	Q93058	tobacco mos	234	26	66.4	505	1	ALAT_SCHRO	P10703	schizosach
162	27	71.1	1616	1	RROO_TMYV	Q93058	tobacco mos	235	26	66.4	505	1	ALAT_SCHRO	P10703	schizosach
163	27	71.1	1616	1	RROO_TMYV	Q93058	tobacco mos	236	26	66.4	505	1	ALAT_SCHRO	P10703	schizosach
164	27	71.1	1616	1	RROO_TMYV	Q93058	tobacco mos	237	26	66.4	505	1	ALAT_SCHRO	P10703	schizosach
165	27	71.1	1616	1	RROO_TMYV	Q93058	tobacco mos	238	26	66.4	505	1	ALAT_SCHRO	P10703	schizosach
166	27	71.1	1616	1	RROO_TMYV	Q93058	tobacco mos	239	26	66.4	505	1	ALAT_SCHRO	P10703	schizosach
167	27	71.1	1616	1	RROO_TMYV	Q93058	tobacco mos	240	26	66.4	505	1	ALAT_SCHRO	P10703	schizosach
168	27	71.1	1616	1	RROO_TMYV	Q93058	tobacco mos	241	26	66.4	505	1	ALAT_SCHRO	P10703	schizosach
169	27	71.1	1616	1	RROO_TMYV	Q93058	tobacco mos	242	26	66.4	505	1	ALAT_SCHRO	P10703	schizosach
170	27	71.1	1616	1	RROO_TMYV	Q93058	tobacco mos	243	26	66.4	505	1	ALAT_SCHRO	P10703	schizosach
171	27	71.1	1616	1	RROO_TMYV	Q93058	tobacco mos	244	26	66.4	505	1	ALAT_SCHRO	P10703	schizosach
172	27	71.1	1616	1	RROO_TMYV	Q93058	tobacco mos	245	26	66.4	505	1	ALAT_SCHRO	P10703	schizosach
173	27	71.1	1616	1	RROO_TMYV	Q93058	tobacco mos	246	26	66.4	505	1	ALAT_SCHRO	P10703	schizosach
174	27	71.1	1616	1	RROO_TMYV	Q93058	tobacco mos	247	26	66.4	505	1	ALAT_SCHRO	P10703	schizosach
175	27	71.1	1616	1	RROO_TMYV	Q93058	tobacco mos	248	26	66.4	505	1	ALAT_SCHRO	P10703	schizosach
176	27	71.1	1616	1	RROO_TMYV	Q93058	tobacco mos	249	26	66.4	505	1	ALAT_SCHRO	P10703	schizosach
177	27	71.1	1616	1	RROO_TMYV	Q93058	tobacco mos	250	26	66.4	505	1	ALAT_SCHRO	P10703	schizosach
178	27	71.1	1616	1	RROO_TMYV	Q93058	tobacco mos	251	26	66.4	505	1	ALAT_SCHRO	P10703	schizosach
179	27	71.1	1616	1	RROO_TMYV	Q93058	tobacco mos	252	26	66.4	505	1	ALAT_SCHRO	P10703	schizosach



253	25	65.8	177	1	APT_MYCPN	P75388 mycoplasma	326	25	65.8	361	1	CG2B_CAEEL	O10653 caenorhabdi
254	25	65.8	178	1	ISPE_CHIMU	O9piv8 chlamydia m	327	25	65.8	363	1	TRNA_PSEAE	O9hv78 pseudomonas
255	25	65.8	180	1	APT_GERCA	O64414 gerbillus c	328	25	65.8	363	1	TRNA_PSEEL	O9rth9 pseudomonas
256	25	65.8	180	1	APT_MASHI	O64427 mastomys hi	329	25	65.8	366	1	HENM_BACSU	P54304 bacillus su
257	25	65.8	180	1	APT_MOUSE	P08030 mus musculu	330	25	65.8	367	1	DRG1_HUMAN	O9y295 homo sapien
258	25	65.8	180	1	APT_MUSPA	P47956 mus pahari	331	25	65.8	367	1	DRG1_MOUSE	P32233 mus musculu
259	25	65.8	180	1	APT_MUSSA	P47957 mus spicile	332	25	65.8	367	1	DRG1_XENLA	P43690 xenopus lae
260	25	65.8	180	1	APT_RAT	P36972 rattus norv	333	25	65.8	368	1	128U_DROME	P32234 drosophila
261	25	65.8	180	1	APT_STOLO	P47958 stochomys l	334	25	65.8	371	1	REC8_BACHD	O9rc99 bacillus ha
262	25	65.8	181	1	APT_DROPS	P54363 drosophila	335	25	65.8	371	1	RSPD_STRMU	O33662 streptococc
263	25	65.8	191	1	GVE1_HALNI	P13044 halobacteri	336	25	65.8	374	1	SPDP_HUMAN	O43791 homo sapien
264	25	65.8	198	1	RS5_ARCPU	O28374 archaeoglob	337	25	65.8	375	1	VATC_ARATH	O9sd57 arabidopsis
265	25	65.8	199	1	COAE_CLOPE	O8x160 clostridium	338	25	65.8	376	1	PTFA_SALTY	P17127 s pls syste
266	25	65.8	201	1	YF04_ARCFU	O28768 archaeoglob	339	25	65.8	386	1	SYM_PYRHO	O95884 pyrococcus
267	25	65.8	205	1	KGUA_CHLTR	O84033 chlamydia t	340	25	65.8	387	1	Y299_METJA	O8xv63 ralsonia s
268	25	65.8	213	1	RIB7_SURSO	P95872 sulfolobus	341	25	65.8	389	1	DC11_RALSO	O57747 methanococ
269	25	65.8	213	1	TPYR_CHIPN	O92698 chlamydia p	342	25	65.8	393	1	CYS3_YEAST	P31373 saccharomyc
270	25	65.8	223	1	IF6_SULSO	O98090 sulfolobus	343	25	65.8	398	1	AROC_AQOAE	O68493 aquifex ae
271	25	65.8	225	1	ALKD_PSEPU	P00885 pseudomonas	344	25	65.8	399	1	SUT3_HUMAN	O75486 homo sapien
272	25	65.8	233	1	RR2_CYPAP	P48132 cyanophora	345	25	65.8	400	1	ACKA_CLOTS	O9y331 clostridium
273	25	65.8	234	1	SL20_YEAST	P39931 saccharomyc	346	25	65.8	401	1	CASP_MOUSE	O9cyd3 mus musculu
274	25	65.8	238	1	RNH2_THEMA	O9x017 thermotoga	347	25	65.8	401	1	CASP_HUMAN	O75718 homo sapien
275	25	65.8	243	1	CY1_EUGGR	P20114 euglena gra	348	25	65.8	402	1	NOR_FUSOX	P23295 fusarium ox
276	25	65.8	247	1	COX2_KLUDA	P20387 kluyveromyc	349	25	65.8	405	1	VG11_HVSA	P24914 herpesvirus
277	25	65.8	247	1	COX2_WILMR	P47918 williopsis	350	25	65.8	411	1	PGK_MYCTU	P23550 escherichia
278	25	65.8	247	1	COX2_WILSA	O23716 arabidopsis	351	25	65.8	412	1	PURK_CORAM	O06821 mycobacteri
279	25	65.8	249	1	PSA3_ARATH	P43376 kluyveromyc	352	25	65.8	414	1	AROA_SUISS	O44678 corynebacte
280	25	65.8	251	1	COX2_KLUTH	P00410 saccharomyc	353	25	65.8	422	1	VT22_CERCA	P27587 ceratitis c
281	25	65.8	251	1	COX2_YEAST	P40766 bacillus su	354	25	65.8	422	1	CISY_ECOLI	P00891 escherichia
282	25	65.8	255	1	YPMR_BACSU	O9pbd0 xyella fas	355	25	65.8	427	1	CISY_SALTY	O68883 salmonella
283	25	65.8	259	1	PEMK_XYLFA	O8019 methanococ	356	25	65.8	427	1	PEPB_ECOS7	P58473 escherichia
284	25	65.8	261	1	E602_METUA	O65962 canine aden	357	25	65.8	427	1	PEPB_ECOS7	P37095 escherichia
285	25	65.8	265	1	E434_ADECC	O96690 canine aden	358	25	65.8	427	1	PEPB_SALTI	O82433 salmonella
286	25	65.8	265	1	E434_ADECC	O06244 mycobacteri	359	25	65.8	427	1	PEPB_SALTY	O9f152 salmonella
287	25	65.8	267	1	CYSO_MYCTU	P14511 shigella fl	360	25	65.8	432	1	PEPB_YERPE	P58475 yersinia pe
288	25	65.8	269	1	33AD_SHIFL	P39610 bacillus su	361	25	65.8	432	1	YRKO_BACSU	P54444 bacillus su
289	25	65.8	271	1	THRD_BACSU	P24172 escherichia	362	25	65.8	441	1	Y640_YEAST	O05990 fonguieria
290	25	65.8	280	1	YIBA_ECOLI	O00164 ictaluriid h	363	25	65.8	448	1	Y640_YEAST	P50079 saccharomyc
291	25	65.8	281	1	VG11_HSVII	O81K9 ralsonia s	364	25	65.8	455	1	CD51_CAEEL	P33439 caenorhabdi
292	25	65.8	285	1	APAH_RALSO	O02760 rhodobacter	365	25	65.8	455	1	RBL_ERYCG	O33438 erythrina c
293	25	65.8	285	1	CY1_RHOSH	O05852 b utp--gluc	366	25	65.8	461	1	Y065_MYCPN	P04070 homo sapien
294	25	65.8	292	1	GTAB_BACSU	O9a817 caulobacter	367	25	65.8	465	1	SAPR_HUMAN	O9nsd7 homo sapien
295	25	65.8	299	1	RDGC_NEIMB	O87408 neisseria g	368	25	65.8	469	1	HOS1_YEAST	O12214 saccharomyc
296	25	65.8	299	1	RDGC_NEIMA	O91v02 neisseria m	369	25	65.8	470	1	PAP1_FOWEY	O9j5b6 towpox vir
297	25	65.8	299	1	RDGC_NEIMB	O91z12 neisseria m	370	25	65.8	472	1	HOS1_YEAST	O9j5b6 towpox vir
298	25	65.8	299	1	RDGC_NEIMB	P39376 escherichia	371	25	65.8	475	1	FAC1_HUMAN	O9y5n0 manduca sex
299	25	65.8	303	1	YJIE_ECOLI	O10298 schizosacch	372	25	65.8	475	1	VATM_MANSE	P31071 escherichia
300	25	65.8	307	1	YD4_SCHPO	O9abey caulobacter	373	25	65.8	486	1	LIPI_YARLI	O99156 yarrowia li
301	25	65.8	308	1	FMT_CAVCR	P52684 klebsiella	374	25	65.8	486	1	LIPI_YARLI	O08429 vibrio angu
302	25	65.8	309	1	MAUR_KLEPN	O979x5 thermoplas	375	25	65.8	486	1	RP54_VIBAN	P17770 catharantu
303	25	65.8	309	1	KHSE_THENO	P47129 saccharomyc	376	25	65.8	500	1	DDC_CATRO	O13397 homo sapien
304	25	65.8	309	1	YJ31_YEAST	P08330 escherichia	377	25	65.8	508	1	KO20_HUMAN	P28937 thogogo vir
305	25	65.8	314	1	KPRS_ECOLI	P15849 salmonella	378	25	65.8	512	1	VENY_THOGV	P38573 mus musculu
306	25	65.8	314	1	KPRS_ECOLI	P42068 bacillus su	379	25	65.8	517	1	HSF2_MOUSE	O12573 candida api
307	25	65.8	319	1	HUTG_BACSU	O39211 arabidopsis	380	25	65.8	519	1	CP5W_CANAP	P13723 avian infec
308	25	65.8	319	1	RP3A_ARATH	P38884 saccharomyc	381	25	65.8	521	1	RREP_TBYK	P70091 oreochromis
309	25	65.8	321	1	YH28_YEAST	P50345 lupinus lut	382	25	65.8	522	1	CPV1_ORENI	P24920 phytophor
310	25	65.8	323	1	RLAO_LUPLU	P19641 escherichia	383	25	65.8	531	1	TRPC_PAYPR	O03933 homo sapien
311	25	65.8	323	1	ISPB_ECOLI	P51954 ratius notv	384	25	65.8	536	1	HSF2_HUMAN	P42833 saccharomyc
312	25	65.8	329	1	CDK7_RAT	P11552 ratius notv	385	25	65.8	544	1	HXTD_YEAST	P21073 vaccinia vi
313	25	65.8	329	1	THRD_SALTY	O59988 synechococ	386	25	65.8	544	1	VAS5_VACCV	P24768 vaccinia vi
314	25	65.8	331	1	KPRS_SYNP7	P40400 bacillus su	387	25	65.8	554	1	VAS5_VACCV	P97318 vaccinia vi
315	25	65.8	332	1	SSUA_BACSU	O55848 synechocyst	388	25	65.8	564	1	DAB1_MOUSE	P07334 schizosacch
316	25	65.8	333	1	KPRS_SYNY3	O67655 aquifex ae	389	25	65.8	593	1	CDRI_SCHPO	O92350 rhizobium m
317	25	65.8	335	1	MDH_AQOAE	P95522 photomidium	390	25	65.8	606	1	EDD_RHIME	O9x0K7 thermotoga
318	25	65.8	338	1	CYE_PHOLA	O03471 penicillium	391	25	65.8	608	1	FLID_THEMA	O62210 mus musculu
319	25	65.8	342	1	ARIS_PENRO	P51953 penicillium	392	25	65.8	612	1	BIR3_MOUSE	P17520 potato leaf
320	25	65.8	344	1	CDK7_CARAU	O03347 m cell divi	393	25	65.8	616	1	RRPO_PLRVL	P11623 potato leaf
321	25	65.8	346	1	CDK7_MOUSE	O14330 schizosacch	394	25	65.8	616	1	RRPO_PLRVL	P56138 heliobacte
322	25	65.8	347	1	KPRS_SCHPO	P31205 euglena gra	395	25	65.8	621	1	GIDA_HELPY	O9j441 neisseria m
323	25	65.8	348	1	CHLI_EUGGR	P20911 xenopus lae	396	25	65.8	628	1	GIDA_NEIMA	O9K190 neisseria m
324	25	65.8	352	1	CDK7_XENLA	O04852 citrobacter	397	25	65.8	628	1	GIDA_NEIMA	
325	25	65.8	355	1	T2C1_CITER		398	25	65.8	628	1	GIDA_NEIMB	

399	25	65.8	629	1	GIDA_XYLFA	09pbu4 xylella fas	472	25	65.8	983	1	CAGE_HELPY	04852 heliobacte
400	25	65.8	629	1	HAP1_RAT	P54756 rattus norv	473	25	65.8	1012	1	IF2C_PHAUV	P57997 phaeosolus v
401	25	65.8	630	1	GIDA_CLOPE	08x131 clostridium	474	25	65.8	1016	1	IF2C_ARATH	09b11 arabidopsis
402	25	65.8	630	1	GIDA_PSEPU	09n039 pseudomonas	475	25	65.8	1018	1	HML1_MYCPN	050365 mycoplasma
403	25	65.8	630	1	GIDA_PSEPU	P25756 pseudomonas	476	25	65.8	1025	1	KG09_HUMAN	09h040 homo sapien
404	25	65.8	634	1	GYR5_BORHE	09z41k1 borrelia he	477	25	65.8	1034	1	BGAL_KLEPN	P06212 klebsiella
405	25	65.8	634	1	NTPI_MCV1	098467 molluscum c	478	25	65.8	1055	1	CC28_SCHPO	010752 schizosacch
406	25	65.8	635	1	PIB2_YEAST	P53191 saccharomyc	479	25	65.8	1067	1	CARB_CLOPE	08xhb3 clostridium
407	25	65.8	639	1	TEPI_ENTFA	P47810 enterococcu	480	25	65.8	1092	1	DP2L_METTH	027578 methanodact
408	25	65.8	639	1	TEPI_ENTFA	P21598 enterococcu	481	25	65.8	1096	1	ATCY_SCHPO	014022 schizosacch
409	25	65.8	639	1	TEPI_ENTFA	O51238 neisseria m	482	25	65.8	1128	1	BEM3_YEAST	P32873 saccharomyc
410	25	65.8	639	1	TEPI_ENTFA	O53770 staphylococ	483	25	65.8	1139	1	DP2L_METFA	094024 methanococ
411	25	65.8	639	1	TEPI_ENTFA	P09797 streptococc	484	25	65.8	1149	1	DBS_MOUSE	O64096 mus musculi
412	25	65.8	639	1	TEPI_ENTFA	O53692 streptococ	485	25	65.8	1150	1	ZACA_HUMAN	012869 homo sapien
413	25	65.8	642	1	PHSA_STRAT	P40032 saccharomyc	486	25	65.8	1164	1	KELI_YEAST	P38883 saccharomyc
414	25	65.8	644	1	YENG_YEAST	O8xuv5 talstonia s	487	25	65.8	1166	1	HERG_HUMAN	008470 drosophila
415	25	65.8	647	1	GIDA_RALSO	084307 chlamydia t	488	25	65.8	1174	1	IF3X_SCHPO	059742 schizosacch
416	25	65.8	650	1	VART_CHLTR	094444 homo sapien	489	25	65.8	1254	1	POLS_EEYV8	P05674 venezuelan
417	25	65.8	657	1	UVRB_HUMAN	P72587 chlamydia p	490	25	65.8	1254	1	POLS_EEYV8	P05674 venezuelan
418	25	65.8	657	1	UVRB_HUMAN	P72587 chlamydia p	491	25	65.8	1254	1	POLS_EEYV8	P05674 venezuelan
419	25	65.8	657	1	UVRB_HUMAN	P72587 chlamydia p	492	25	65.8	1254	1	POLS_EEYV8	P05674 venezuelan
420	25	65.8	657	1	UVRB_HUMAN	P72587 chlamydia p	493	25	65.8	1254	1	POLS_EEYV8	P05674 venezuelan
421	25	65.8	657	1	UVRB_HUMAN	P72587 chlamydia p	494	25	65.8	1254	1	POLS_EEYV8	P05674 venezuelan
422	25	65.8	657	1	UVRB_HUMAN	P72587 chlamydia p	495	25	65.8	1254	1	POLS_EEYV8	P05674 venezuelan
423	25	65.8	657	1	UVRB_HUMAN	P72587 chlamydia p	496	25	65.8	1254	1	POLS_EEYV8	P05674 venezuelan
424	25	65.8	657	1	UVRB_HUMAN	P72587 chlamydia p	497	25	65.8	1254	1	POLS_EEYV8	P05674 venezuelan
425	25	65.8	657	1	UVRB_HUMAN	P72587 chlamydia p	498	25	65.8	1254	1	POLS_EEYV8	P05674 venezuelan
426	25	65.8	657	1	UVRB_HUMAN	P72587 chlamydia p	499	25	65.8	1254	1	POLS_EEYV8	P05674 venezuelan
427	25	65.8	657	1	UVRB_HUMAN	P72587 chlamydia p	500	25	65.8	1254	1	POLS_EEYV8	P05674 venezuelan
428	25	65.8	657	1	UVRB_HUMAN	P72587 chlamydia p	501	25	65.8	1254	1	POLS_EEYV8	P05674 venezuelan
429	25	65.8	657	1	UVRB_HUMAN	P72587 chlamydia p	502	25	65.8	1254	1	POLS_EEYV8	P05674 venezuelan
430	25	65.8	657	1	UVRB_HUMAN	P72587 chlamydia p	503	25	65.8	1254	1	POLS_EEYV8	P05674 venezuelan
431	25	65.8	657	1	UVRB_HUMAN	P72587 chlamydia p	504	25	65.8	1254	1	POLS_EEYV8	P05674 venezuelan
432	25	65.8	657	1	UVRB_HUMAN	P72587 chlamydia p	505	25	65.8	1254	1	POLS_EEYV8	P05674 venezuelan
433	25	65.8	657	1	UVRB_HUMAN	P72587 chlamydia p	506	25	65.8	1254	1	POLS_EEYV8	P05674 venezuelan
434	25	65.8	657	1	UVRB_HUMAN	P72587 chlamydia p	507	25	65.8	1254	1	POLS_EEYV8	P05674 venezuelan
435	25	65.8	657	1	UVRB_HUMAN	P72587 chlamydia p	508	25	65.8	1254	1	POLS_EEYV8	P05674 venezuelan
436	25	65.8	657	1	UVRB_HUMAN	P72587 chlamydia p	509	25	65.8	1254	1	POLS_EEYV8	P05674 venezuelan
437	25	65.8	657	1	UVRB_HUMAN	P72587 chlamydia p	510	25	65.8	1254	1	POLS_EEYV8	P05674 venezuelan
438	25	65.8	657	1	UVRB_HUMAN	P72587 chlamydia p	511	25	65.8	1254	1	POLS_EEYV8	P05674 venezuelan
439	25	65.8	657	1	UVRB_HUMAN	P72587 chlamydia p	512	25	65.8	1254	1	POLS_EEYV8	P05674 venezuelan
440	25	65.8	657	1	UVRB_HUMAN	P72587 chlamydia p	513	25	65.8	1254	1	POLS_EEYV8	P05674 venezuelan
441	25	65.8	657	1	UVRB_HUMAN	P72587 chlamydia p	514	25	65.8	1254	1	POLS_EEYV8	P05674 venezuelan
442	25	65.8	657	1	UVRB_HUMAN	P72587 chlamydia p	515	25	65.8	1254	1	POLS_EEYV8	P05674 venezuelan
443	25	65.8	657	1	UVRB_HUMAN	P72587 chlamydia p	516	25	65.8	1254	1	POLS_EEYV8	P05674 venezuelan
444	25	65.8	657	1	UVRB_HUMAN	P72587 chlamydia p	517	25	65.8	1254	1	POLS_EEYV8	P05674 venezuelan
445	25	65.8	657	1	UVRB_HUMAN	P72587 chlamydia p	518	25	65.8	1254	1	POLS_EEYV8	P05674 venezuelan
446	25	65.8	657	1	UVRB_HUMAN	P72587 chlamydia p	519	25	65.8	1254	1	POLS_EEYV8	P05674 venezuelan
447	25	65.8	657	1	UVRB_HUMAN	P72587 chlamydia p	520	25	65.8	1254	1	POLS_EEYV8	P05674 venezuelan
448	25	65.8	657	1	UVRB_HUMAN	P72587 chlamydia p	521	25	65.8	1254	1	POLS_EEYV8	P05674 venezuelan
449	25	65.8	657	1	UVRB_HUMAN	P72587 chlamydia p	522	25	65.8	1254	1	POLS_EEYV8	P05674 venezuelan
450	25	65.8	657	1	UVRB_HUMAN	P72587 chlamydia p	523	25	65.8	1254	1	POLS_EEYV8	P05674 venezuelan
451	25	65.8	657	1	UVRB_HUMAN	P72587 chlamydia p	524	25	65.8	1254	1	POLS_EEYV8	P05674 venezuelan
452	25	65.8	657	1	UVRB_HUMAN	P72587 chlamydia p	525	25	65.8	1254	1	POLS_EEYV8	P05674 venezuelan
453	25	65.8	657	1	UVRB_HUMAN	P72587 chlamydia p	526	25	65.8	1254	1	POLS_EEYV8	P05674 venezuelan
454	25	65.8	657	1	UVRB_HUMAN	P72587 chlamydia p	527	25	65.8	1254	1	POLS_EEYV8	P05674 venezuelan
455	25	65.8	657	1	UVRB_HUMAN	P72587 chlamydia p	528	25	65.8	1254	1	POLS_EEYV8	P05674 venezuelan
456	25	65.8	657	1	UVRB_HUMAN	P72587 chlamydia p	529	25	65.8	1254	1	POLS_EEYV8	P05674 venezuelan
457	25	65.8	657	1	UVRB_HUMAN	P72587 chlamydia p	530	25	65.8	1254	1	POLS_EEYV8	P05674 venezuelan
458	25	65.8	657	1	UVRB_HUMAN	P72587 chlamydia p	531	25	65.8	1254	1	POLS_EEYV8	P05674 venezuelan
459	25	65.8	657	1	UVRB_HUMAN	P72587 chlamydia p	532	25	65.8	1254	1	POLS_EEYV8	P05674 venezuelan
460	25	65.8	657	1	UVRB_HUMAN	P72587 chlamydia p	533	25	65.8	1254	1	POLS_EEYV8	P05674 venezuelan
461	25	65.8	657	1	UVRB_HUMAN	P72587 chlamydia p	534	25	65.8	1254	1	POLS_EEYV8	P05674 venezuelan
462	25	65.8	657	1	UVRB_HUMAN	P72587 chlamydia p	535	25	65.8	1254	1	POLS_EEYV8	P05674 venezuelan
463	25	65.8	657	1	UVRB_HUMAN	P72587 chlamydia p	536	25	65.8	1254	1	POLS_EEYV8	P05674 venezuelan
464	25	65.8	657	1	UVRB_HUMAN	P72587 chlamydia p	537	25	65.8	1254	1	POLS_EEYV8	P05674 venezuelan
465	25	65.8	657	1	UVRB_HUMAN	P72587 chlamydia p	538	25	65.8	1254	1	POLS_EEYV8	P05674 venezuelan
466	25	65.8	657	1	UVRB_HUMAN	P72587 chlamydia p	539	25	65.8	1254	1	POLS_EEYV8	P05674 venezuelan
467	25	65.8	657	1	UVRB_HUMAN	P72587 chlamydia p	540	25	65.8	1254	1	POLS_EEYV8	P05674 venezuelan
468	25	65.8	657	1	UVRB_HUMAN	P72587 chlamydia p	541	25	65.8	1254	1	POLS_EEYV8	P05674 venezuelan
469	25	65.8	657	1	UVRB_HUMAN	P72587 chlamydia p	542	25	65.8	1254	1	POLS_EEYV8	P05674 venezuelan
470	25	65.8	657	1	UVRB_HUMAN	P72587 chlamydia p	543	25	65.8	1254	1	POLS_EEYV8	P05674 venezuelan
471	25	65.8	657	1	UVRB_HUMAN	P72587 chlamydia p	544	25	65.8	1254	1	POLS_EEYV8	P05674 venezuelan

545	24	63.2	122	1	YPC3_CLAPU	P22371 claviceps p	618	24	63.2	284	1	TPM1_RAT	P04692 ratus norv
546	24	63.2	123	1	RL7_NEIPE	Q915m1 neisseria p	619	24	63.2	284	1	TPM1_XENIA	Q01173 xenopus lae
547	24	63.2	126	1	LCA_ORMAN	P30805 ornithohym	620	24	63.2	299	1	CAC4_STREP	P58133 streptococ
548	24	63.2	136	1	DEF_CLOHE	O08450 clostridium	621	24	63.2	299	1	HAC2_STREY	P58098 streptococ
549	24	63.2	140	1	PUCU_ECOLI	P1555 escherichia	622	24	63.2	301	1	CAPB_BOVIN	P79136 bos taurus
550	24	63.2	142	1	GLMG_HUMAN	O60234 homo sapien	623	24	63.2	302	1	RDGC_XYLFA	O9f19 xyella fas
551	24	63.2	143	1	FER_WHEAT	P00228 triticum ae	624	24	63.2	304	1	HAC1_STREY	O54713 streptococ
552	24	63.2	146	1	BAHG_VITST	P04252 vitreoscill	625	24	63.2	306	1	CAC3_STREN	O54800 streptococ
553	24	63.2	147	1	PER_SILPR	P4659 silene prat	626	24	63.2	307	1	NAHH_PSEBU	P08127 pseudomonas
554	24	63.2	148	1	GLP3_GLYDI	P21660 glycera dib	627	24	63.2	308	1	BGL2_CANAL	P43070 candida alb
555	24	63.2	149	1	URRE_BACSB	Q07401 bacillus sp	628	24	63.2	311	1	YCH4_YEAST	P25600 saccharomyc
556	24	63.2	150	1	RBFA_BRIME	O8yeb4 bruceella me	629	24	63.2	311	1	PYRB_LACPL	P77883 lactobacill
557	24	63.2	157	1	RS7_EIKCO	P35642 eikenella c	630	24	63.2	315	1	KPRS_BACUL	P42816 bacillus ca
558	24	63.2	162	1	TERM_BLP7	P16937 bacterioph	631	24	63.2	315	1	KPRS_BUCAL	P57266 buchnera ap
559	24	63.2	162	1	YX11_BACSU	P42301 haemophilus su	632	24	63.2	315	1	KPRS_HAETN	P44328 haemophilus
560	24	63.2	163	1	PGPA_HAEIN	O9kvu3 vibrio chol	633	24	63.2	317	1	KPRS_BACSU	P14193 bacillus su
561	24	63.2	169	1	DEF_VIBCH	O9kvu3 vibrio chol	634	24	63.2	318	1	KPRS_HELPJ	O92141 helicobacte
562	24	63.2	169	1	SUUA_ECOLI	O8yeb4 bruceella me	635	24	63.2	318	1	KPRS_HELPJ	P61844 helicobacte
563	24	63.2	171	1	LACB_LACLA	P23495 lactococcus	636	24	63.2	319	1	HPRK_TREPA	O83600 treponema p
564	24	63.2	178	1	HPRT_AOUAE	O66821 aquilex aeo	637	24	63.2	324	1	OTCA_BACSU	P18186 bacillus su
565	24	63.2	179	1	APT_HUMAN	P07741 homo sapien	638	24	63.2	324	1	TKRA_ECO57	P18220 escherichia
566	24	63.2	180	1	APT_CRILO	P47952 citreulius	639	24	63.2	324	1	TKRA_ECO57	P37666 escherichia
567	24	63.2	182	1	APT_DROME	P12426 dirosophila	640	24	63.2	325	1	BET2_YEAST	P20133 saccharomyc
568	24	63.2	184	1	MLRL_SCHPO	O9uuv9 schizosacch	641	24	63.2	326	1	ATP3_IPOBA	P26360 ipomeea bat
569	24	63.2	185	1	RIMM_RALSO	O8yov9 ralstonia s	642	24	63.2	329	1	ISPA_RHISN	O06844 halobacteri
570	24	63.2	185	1	APT_CORGL	O87330 corynebacte	643	24	63.2	330	1	GGPP_SULAC	P39464 sulfolobus
571	24	63.2	187	1	ORN_NEIMA	O91w09 neisseria m	644	24	63.2	330	1	ISPA_RHISN	P55539 rhizobium s
572	24	63.2	187	1	ORN_NEIMA	O91xw1 neisseria m	645	24	63.2	332	1	CAPT_STYAU	P39838 staphylococ
573	24	63.2	192	1	DYR_CANAL	P22906 candida alb	646	24	63.2	334	1	HOLA_ECOLI	P26630 escherichia
574	24	63.2	195	1	NIR0_AZOVI	P11068 azotobacter	647	24	63.2	349	1	DIA2_MOUSE	O70566 mus musculu
575	24	63.2	203	1	Y461_PYROHO	O58814 pyrococcus	648	24	63.2	349	1	YAAE_SCHPO	O09732 schizosacch
576	24	63.2	209	1	IPVR_CHILMU	O98f11 chlamydia m	649	24	63.2	350	1	EUTR_ECOLI	P36547 escherichia
577	24	63.2	209	1	IPVR_CHILTR	O84777 chlamydia t	650	24	63.2	350	1	EUTR_SALTY	O9zf57 salmonella
578	24	63.2	213	1	EF1X_CABEL	P34460 caenorhabdi	651	24	63.2	350	1	DPNF_BACSU	P58639 methanopyru
579	24	63.2	213	1	PYRE_BUCAL	P57622 buchnera ap	652	24	63.2	353	1	DPNF_BACSU	O05252 bacillus su
580	24	63.2	218	1	ESM2_DROME	O97177 dirosophila	653	24	63.2	353	1	DPNF_ARATH	O42546 arabidopsis
581	24	63.2	226	1	TPIS_METKA	O8ut09 methanopyru	654	24	63.2	355	1	C3X1_HUMAN	P49238 homo sapien
582	24	63.2	229	1	PCRB_BACHD	O9Kf39 bacillus ha	655	24	63.2	362	1	LEU3_PICAN	P34723 plichia angu
583	24	63.2	230	1	HGXR_TOXGO	O26597 toxoplasma	656	24	63.2	368	1	SERC_NEIMA	O34370 neisseria m
584	24	63.2	237	1	PYRE_RALSO	O8y419 ralstonia s	657	24	63.2	368	1	SERC_NEIMA	P57070 neisseria m
585	24	63.2	238	1	KAD2_HUMAN	P54819 homo sapien	658	24	63.2	369	1	TRAI_LACHE	P35680 lactobacill
586	24	63.2	240	1	RS4_CHLS6	P08166 bos taurus	659	24	63.2	370	1	RECF_BACSU	P05651 bacillus su
587	24	63.2	242	1	YU68_CABEL	Q40941 chlorarachn	660	24	63.2	372	1	3BH1_MESAU	O60555 m 3 beta-hy
588	24	63.2	244	1	SUMT_PSEFL	P52057 caenorhabdi	661	24	63.2	372	1	3BH2_MESAU	P24615 m 3 beta-hy
589	24	63.2	247	1	COX2_BRENA	P33725 pseudomomy	662	24	63.2	372	1	3BH2_MESAU	O64421 m 3 beta-hy
590	24	63.2	248	1	LICA_MYCHO	P43052 mycoplasma	663	24	63.2	372	1	3BH3_MESAU	O53296 mesocricetu
591	24	63.2	249	1	Y7SR_METFE	P21336 methanother	664	24	63.2	372	1	3BH5_MOUSE	O61767 mus musculu
592	24	63.2	251	1	THIG_ERWAM	Q91812 ervania amy	665	24	63.2	372	1	3BH5_MOUSE	O61694 mus musculu
593	24	63.2	252	1	UPPS_CHLTR	O84456 chlamydia t	666	24	63.2	374	1	SPB8_HUMAN	P50452 homo sapien
594	24	63.2	253	1	DAPB_HELPY	P94844 helicobacte	667	24	63.2	379	1	DEFA_LYCES	P51107 lycopersico
595	24	63.2	254	1	YD53_MYCNU	O11023 mycobacteri	668	24	63.2	380	1	DEFA_PENHY	P14720 petunia hyb
596	24	63.2	261	1	CBIT_SALTY	O05591 salmonella	669	24	63.2	380	1	TGT_STREY	O974h1 streptococ
597	24	63.2	263	1	EUTL_ECOLI	P76554 escherichia	670	24	63.2	380	1	TGT_STREY	O9a116 streptococ
598	24	63.2	267	1	EUTL_ECOLI	O9zf54 salmonella	671	24	63.2	382	1	TGT_LACLA	O9c154 lactococcus
599	24	63.2	267	1	THIG_AOUAE	O67926 aquilex aeo	672	24	63.2	386	1	LY2_ADECC	O43066 pisin sativ
600	24	63.2	267	1	THIG_AOUAE	P46647 cercocobus	673	24	63.2	389	1	SPSC_BACSU	O65948 canine aden
601	24	63.2	271	1	ILIA_CERTO	P42407 bacillus*su	674	24	63.2	389	1	LY2_ADECC	P39623 bacillus su
602	24	63.2	272	1	YVDA_BACSU	P17548 pseudomomy	675	24	63.2	389	1	LY2_ADECC	O95330 oryctolagus
603	24	63.2	277	1	BPBD_PSEST	P38878 saccharomyc	676	24	63.2	391	1	LY2_ADECC	O43064 pisin sativ
604	24	63.2	278	1	GAU1_PSEAE	O91291 pseudomomy	677	24	63.2	391	1	LY2_ADECC	P22857 streptomyce
605	24	63.2	279	1	GAU2_PSEAE	O99633 pseudomomy	678	24	63.2	393	1	LY2_ADECC	O03841 rhizobium m
606	24	63.2	280	1	PARB_CHLMU	O99633 pseudomomy	679	24	63.2	394	1	FLAI_RHIME	O03842 rhizobium m
607	24	63.2	281	1	UGPE_ECOLI	P10906 escherichia	680	24	63.2	397	1	FLAI_RHIME	O26662 methanobact
608	24	63.2	281	1	YGBH_CAEEL	Q9u2c3 caenorhabdi	681	24	63.2	397	1	FLAI_RHIME	P16791 human cytom
609	24	63.2	284	1	TPM1_BRARE	P13104 brachydanio	682	24	63.2	401	1	U5UO_HCMVA	O92367 rickettsia
610	24	63.2	284	1	TPM1_CHICK	P04268 gallus gall	683	24	63.2	402	1	U5UO_HCMVA	P29452 mus musculu
611	24	63.2	284	1	TPM1_COTTA	P58773 coturnix co	684	24	63.2	403	1	YDE4_SCHRO	P43527 ratus norv
612	24	63.2	284	1	TPM1_COTTA	P09493 homo sapien	685	24	63.2	404	1	YDE4_SCHRO	O10438 schizosacch
613	24	63.2	284	1	TPM1_HUMAN	P58771 mus musculu	686	24	63.2	404	1	YDE4_SCHRO	O00441 saccharopol
614	24	63.2	284	1	TPM1_MOUSE	P42639 sus scrofa	687	24	63.2	405	1	CPXJ_SACBR	Q9h211 sus scrofa
615	24	63.2	284	1	TPM1_PIG	P58772 oryctolagus	688	24	63.2	405	1	CPXJ_SACBR	P33271 saccharopol
616	24	63.2	284	1	TPM1_RABIT	P13105 rana tempor	689	24	63.2	406	1	ILBC_HORSE	O9tlv3 equus cabal
617	24	63.2	284	1	TPM1_RABIT	P13105 rana tempor	690	24	63.2	406	1	VPS_PSIINU	O9slx9 psittotum nu

691	24	63.2	409	1	HMDH_SUIISO	008424	sulfolobus	764	24	63.2	548	1	SYN_BRUMA	P10723	brugia mala
692	24	63.2	410	1	PEK_AERPE	Q9Y57	aeropyrum p	765	24	63.2	550	1	SOAI_CERAE	Q77760	cercoptihac
693	24	63.2	410	1	LIJ8_CABEL	P34371	caenorhabd	766	24	63.2	550	1	SOAI_HUMAN	P35510	homo sapien
694	24	63.2	413	1	DCHS_LYCES	P54772	lycopestico	767	24	63.2	550	1	SOAI_HUMAN	Q77761	maeca fasc
695	24	63.2	415	1	KAPR_YEAST	P07272	saccharomyc	768	24	63.2	557	1	PKSJ_MACFA	P40806	baclillus su
696	24	63.2	416	1	G3PA_GRAVE	P30724	gracilaria	769	24	63.2	558	1	EXO2_BP69	P04522	bacterioph
697	24	63.2	418	1	V006_FOMPV	P14361	fowipox vir	770	24	63.2	560	1	EXO2_BP14	Q92851	chlamydia p
698	24	63.2	426	1	CLIPX_MCTU	O53184	mycobacteri	771	24	63.2	568	1	SYN_CHLPN	Q92851	chlamydia p
699	24	63.2	426	1	NUOF_AOUAE	O66841	aquifex aeo	772	24	63.2	574	1	PPNK_METJA	O58827	methanococ
700	24	63.2	428	1	TIG_CLOPE	O8XK0	clostridium	773	24	63.2	587	1	SYT3_MOUSE	O35881	mus musculu
701	24	63.2	428	1	V234_FOMPV	P14368	fowipox vir	774	24	63.2	588	1	SYT3_MOUSE	O35881	mus musculu
702	24	63.2	430	1	C139_MYCTU	O86330	mycobacteri	775	24	63.2	590	1	US22_HOMVA	Q90941	homo sapien
703	24	63.2	435	1	ASPG_MERTH	O26802	methanobact	776	24	63.2	593	1	US22_HOMVA	Q90941	homo sapien
704	24	63.2	435	1	CKI2_SCHPO	P40234	schizosacch	777	24	63.2	599	1	YAOB_SCHPO	O10712	homo sapien
705	24	63.2	435	1	HEM1_LISIN	Q928F7	listeria in	778	24	63.2	600	1	GIDA_DEIRA	O10110	schizosacch
706	24	63.2	436	1	RBL_DROBR	P28404	drosophila bur	779	24	63.2	604	1	VEI_HPV38	O02043	delnecococ
707	24	63.2	436	1	YE09_HAEIN	P44183	haemophilus	779	24	63.2	604	1	VEI_HPV38	O02043	delnecococ
708	24	63.2	442	1	LICH_BACSU	P28410	drosophila pet	780	24	63.2	605	1	VEI_HPV38	O02043	delnecococ
709	24	63.2	442	1	LICH_BACSU	P46320	baclillus su	781	24	63.2	604	1	VEI_HPV38	O02043	delnecococ
710	24	63.2	445	1	MO50_CABEL	O17919	caenorhabd	782	24	63.2	606	1	ALB1_XENLA	O08090	human papil
711	24	63.2	447	1	RRA_FUGRU	Q9W523	fugu rubrip	783	24	63.2	606	1	ALB1_XENLA	O08090	human papil
712	24	63.2	448	1	Y260_SYNX3	P54716	baclillus su	784	24	63.2	610	1	ALB2_XENLA	O05111	human papil
713	24	63.2	449	1	AGAL_BACSU	O54716	baclillus su	785	24	63.2	611	1	VATA_TPYCO	O08759	xenopus lae
714	24	63.2	451	1	AGAL_BACSU	O54716	baclillus su	786	24	63.2	611	1	GIDA_MYCPU	P14872	xenopus lae
715	24	63.2	451	1	GLUS_SALTU	P30877	salmonella	787	24	63.2	619	1	GIDA_CAMUE	O26975	trypanobact
716	24	63.2	451	1	GLUS_SALTU	P30877	salmonella	788	24	63.2	621	1	GIDA_CAMUE	O26975	trypanobact
717	24	63.2	452	1	TREC_HELPU	P14018	coturnix co	789	24	63.2	621	1	GIDA_CAMUE	O26975	trypanobact
718	24	63.2	452	1	TREC_HELPU	P14018	coturnix co	790	24	63.2	621	1	GIDA_CAMUE	O26975	trypanobact
719	24	63.2	452	1	YOX4_CABEL	Q92JUH	helicobacte	791	24	63.2	625	1	GNX3_RAT	O09478	mycoplasma
720	24	63.2	453	1	MALE_PYRAB	O25867	helicobacte	792	24	63.2	625	1	GNX3_RAT	O09478	mycoplasma
721	24	63.2	457	1	MORE_BACSU	O9Y297	pyrococcus	793	24	63.2	626	1	GIDA_STAM	O09478	mycoplasma
722	24	63.2	458	1	ME31_DROME	P66113	baclillus su	794	24	63.2	627	1	GIDA_STAM	O09478	mycoplasma
723	24	63.2	466	1	RBI_DROBT	P23128	drosophila	795	24	63.2	628	1	GNX3_RAT	O09478	mycoplasma
724	24	63.2	467	1	GLMA_AZOVI	P28403	drosophila bin	796	24	63.2	629	1	GIDA_STAM	O09478	mycoplasma
725	24	63.2	470	1	GLMA_AZOVI	P28403	drosophila bin	797	24	63.2	629	1	GIDA_STAM	O09478	mycoplasma
726	24	63.2	471	1	HR52_YEAST	Q04515	baclillus su	798	24	63.2	629	1	GIDA_STAM	O09478	mycoplasma
727	24	63.2	471	1	HR52_YEAST	Q04515	baclillus su	799	24	63.2	630	1	GIDA_STAM	O09478	mycoplasma
728	24	63.2	471	1	HR52_YEAST	Q04515	baclillus su	800	24	63.2	630	1	GIDA_STAM	O09478	mycoplasma
729	24	63.2	474	1	SE1A_YARLI	P18979	yarrowia li	801	24	63.2	632	1	GIDA_STAM	O09478	mycoplasma
730	24	63.2	477	1	LACG_LACCA	P14696	lactobacilli	802	24	63.2	636	1	GIDA_STAM	O09478	mycoplasma
731	24	63.2	478	1	MP22_ECOLI	P05101	escherichia	803	24	63.2	637	1	GIDA_STAM	O09478	mycoplasma
732	24	63.2	482	1	ASPA_ECOLI	P04422	escherichia	804	24	63.2	641	1	GIDA_STAM	O09478	mycoplasma
733	24	63.2	484	1	ASPA_ECOLI	P04422	escherichia	805	24	63.2	644	1	GIDA_STAM	O09478	mycoplasma
734	24	63.2	484	1	GATB_THEMA	P33109	seriatia ma	806	24	63.2	644	1	GIDA_STAM	O09478	mycoplasma
735	24	63.2	485	1	HOXA_BRAJA	O9X100	thermotoga	807	24	63.2	645	1	GIDA_STAM	O09478	mycoplasma
736	24	63.2	486	1	PODX_RAT	P31908	bradyrhizob	808	24	63.2	645	1	GIDA_STAM	O09478	mycoplasma
737	24	63.2	489	1	RTG3_YEAST	Q9WT2	rattus norv	809	24	63.2	650	1	GIDA_STAM	O09478	mycoplasma
738	24	63.2	490	1	PNBA_BACSU	P37967	baclillus su	810	24	63.2	653	1	GIDA_STAM	O09478	mycoplasma
739	24	63.2	490	1	V035_CLOPE	Q06373	clostridium	811	24	63.2	653	1	GIDA_STAM	O09478	mycoplasma
740	24	63.2	493	1	V888_MYCTU	Q10549	mycobacteri	812	24	63.2	662	1	GIDA_STAM	O09478	mycoplasma
741	24	63.2	493	1	DCMG_ECOLI	O05822	homo sapien	813	24	63.2	663	1	GIDA_STAM	O09478	mycoplasma
742	24	63.2	494	1	RBSA_HAEIN	P44735	haemophilus	814	24	63.2	663	1	GIDA_STAM	O09478	mycoplasma
743	24	63.2	494	1	VATB_PLAFA	O25691	plasmodium	815	24	63.2	679	1	GIDA_STAM	O09478	mycoplasma
744	24	63.2	496	1	VGA_BPAL3	P25243	bacterioph	816	24	63.2	682	1	GIDA_STAM	O09478	mycoplasma
745	24	63.2	496	1	AMVB_MEDSA	P25244	bacterioph	817	24	63.2	694	1	GIDA_STAM	O09478	mycoplasma
746	24	63.2	498	1	PRIM_AOUAE	O22385	medicago sa	818	24	63.2	694	1	GIDA_STAM	O09478	mycoplasma
747	24	63.2	499	1	AMPA_CHLTR	O67465	aquifex aeo	819	24	63.2	697	1	GIDA_STAM	O09478	mycoplasma
748	24	63.2	499	1	AMPA_CHLTR	P38019	chlamydia m	820	24	63.2	697	1	GIDA_STAM	O09478	mycoplasma
749	24	63.2	503	1	PODX_MOUSE	O84049	chlamydia t	821	24	63.2	701	1	GIDA_STAM	O09478	mycoplasma
750	24	63.2	508	1	YBCK_ECOLI	O97098	mus musculu	822	24	63.2	705	1	GIDA_STAM	O09478	mycoplasma
751	24	63.2	512	1	YR33_MYCTU	P77698	escherichia	823	24	63.2	712	1	GIDA_STAM	O09478	mycoplasma
752	24	63.2	514	1	MALF_ECOLI	O33238	mycobacteri	824	24	63.2	713	1	GIDA_STAM	O09478	mycoplasma
753	24	63.2	514	1	MALF_ECOLI	O33238	mycobacteri	825	24	63.2	713	1	GIDA_STAM	O09478	mycoplasma
754	24	63.2	514	1	MALF_ECOLI	O33238	mycobacteri	826	24	63.2	714	1	GIDA_STAM	O09478	mycoplasma
755	24	63.2	519	1	CP5V_CANAP	P24667	salmonella	827	24	63.2	714	1	GIDA_STAM	O09478	mycoplasma
756	24	63.2	519	1	CP5V_CANAP	P24667	salmonella	828	24	63.2	724	1	GIDA_STAM	O09478	mycoplasma
757	24	63.2	522	1	RHO_YREPA	O83281	treponema p	829	24	63.2	729	1	GIDA_STAM	O09478	mycoplasma
758	24	63.2	530	1	TP68_SULSH	P43113	saccharomyc	830	24	63.2	736	1	GIDA_STAM	O09478	mycoplasma
759	24	63.2	535	1	TP68_SULSH	O05207	sulfolobus	831	24	63.2	736	1	GIDA_STAM	O09478	mycoplasma
760	24	63.2	537	1	ARSB_FELCA	Q972F0	sulfolobus	832	24	63.2	743	1	GIDA_STAM	O09478	mycoplasma
761	24	63.2	537	1	ARSB_FELCA	P33727	fells silve	833	24	63.2	743	1	GIDA_STAM	O09478	mycoplasma
762	24	63.2	540	1	AXR1_ARATH	P24507	discopyge o	834	24	63.2	756	1	GIDA_STAM	O09478	mycoplasma
763	24	63.2	543	1	TCPI_HUMAN	P42744	arabidopsis	835	24	63.2	773	1	GIDA_STAM	O09478	mycoplasma
764	24	63.2	546	1	Y220_HUMAN	O93632	homo sapien	836	24	63.2	782	1	GIDA_STAM	O09478	mycoplasma
765	24	63.2	546	1	Y220_HUMAN	O93632	homo sapien	836	24	63.2	815	1	GIDA_STAM	O09478	mycoplasma

837	24	63.2	837	1	SVV_AERPE	O9yaz0 aetopyrum p	910	23	60.5	95	1	Y104_ADE07	P05666 human adeno
838	24	63.2	838	1	RNI9_HUMAN	O9nv58 homo sapien	911	23	60.5	97	1	FER1_APHFL	P00244 aphaniizomen
839	24	63.2	840	1	RNI9_MOUSE	P50636 mus musculu	912	23	60.5	99	1	YB87_MERTH	O27255 methanobact
840	24	63.2	853	1	DI43_HUMAN	O9ns46 homo sapien	913	23	60.5	100	1	NOLM_PHTIN	O37598 phylophthor
841	24	63.2	859	1	RP42_METJA	P53246 methanococc	914	23	60.5	106	1	Y12K_BP4	P10278 bacteriopho
842	24	63.2	865	1	YG2D_YEAST	P33246 saccharomyc	915	23	60.5	107	1	RLA1_CHLRE	P29763 chlamydomon
843	24	63.2	865	1	FLUG_EMENI	P38094 escherichia	916	23	60.5	109	1	WMT1_IACAO	P26147 influenza a
844	24	63.2	886	1	SM6A_MOUSE	O35464 mus musculu	917	23	60.5	112	1	PER_PSEAE	O51383 pseudomonas
845	24	63.2	896	1	IF2_XIFOX	O9zf28 klebsiella	918	23	60.5	113	1	SU13_HUMAN	O9un99 homo sapien
846	24	63.2	897	1	IF2_ENTCL	O9zf25 enterobacte	919	23	60.5	113	1	SU11_MOUSE	O60739 homo sapien
847	24	63.2	897	1	SY1_FERPE	O9xb44 feravidobact	920	23	60.5	113	1	SU11_MOUSE	P41567 homo sapien
848	24	63.2	904	1	TLR3_HUMAN	O15455 homo sapien	921	23	60.5	113	1	RBA_A_STAM	P48024 mus musculu
849	24	63.2	905	1	VP3_AHSV4	P32509 african hor	922	23	60.5	116	1	YRF4_SHIFL	O994K2 staphylococ
850	24	63.2	905	1	NIOG_ECOLI	P33602 escherichia	923	23	60.5	119	1	PAAD_RALSO	P37790 shigella fl
851	24	63.2	907	1	DPOL_BORBU	O51498 borrelia bu	924	23	60.5	120	1	CVS1_SCHPO	O8xv46 ralsionia s
852	24	63.2	911	1	SECA_STRGR	P46835 mycobacteri	925	23	60.5	124	1	FAEL_ANOPU	O94391 schizosacch
853	24	63.2	911	1	DPOL_MYCLB	P95759 streptomyc	926	23	60.5	125	1	Y568_METTA	O90239 anolis pulc
854	24	63.2	940	1	RHG4_HUMAN	P96171 homo sapien	927	23	60.5	127	1	UCR7_KLIDA	O57988 methanococ
855	24	63.2	946	1	MIC1_YEAST	P53258 saccharomyc	928	23	60.5	130	1	LYC2_CANFA	P58858 sulfolobus
856	24	63.2	950	1	KEMS_HUMAN	P07333 homo sapien	929	23	60.5	130	1	LYCK_SHEEP	P49345 kluyveromyc
857	24	63.2	972	1	LONH_PYPAB	O11208 sarcophaga	930	23	60.5	130	1	LYC2_CANFA	P81709 canis famli
858	24	63.2	996	1	LONH_PYPAB	O9uyvc6 pyrococcus	931	23	60.5	130	1	LYC_RABIT	P16973 oryctolagus
859	24	63.2	998	1	MEF_HELPY	O26066 helicobacte	932	23	60.5	132	1	Y16K_SSVI	P20195 sulfolobus
860	24	63.2	1001	1	MEF_HELPY	O9zj57 vibrio chol	933	23	60.5	133	1	HNS_PROVU	P18818 proteus vul
861	24	63.2	1013	1	SBCC_VIBCH	O9zj57 vibrio chol	934	23	60.5	134	1	YBGC_ECOLI	P08999 escherichia
862	24	63.2	1013	1	CYPE_BACSU	O08336 bacillus su	935	23	60.5	137	1	LYC_BOMO	P48816 bombyx mori
863	24	63.2	1071	1	UBP7_YEAST	P40453 saccharomyc	936	23	60.5	137	1	RBA_A_AGR15	O8uf52 agrobacteri
864	24	63.2	1085	1	CUT7_SCHPO	P24339 schizosacch	937	23	60.5	140	1	VP_C_BPMU	P06052 bacteriopho
865	24	63.2	1085	1	DIA_DROME	P48608 drosophila	938	23	60.5	141	1	LYCN_BOVIN	P80189 bos taurus
866	24	63.2	1099	1	MYE_MOUSE	P70248 mus musculu	939	23	60.5	147	1	YDFE_SCHPO	O10448 schizosacch
867	24	63.2	1101	1	DI42_HUMAN	O60879 homo sapien	940	23	60.5	148	1	LYC1_RAT	P00697 rattus norv
868	24	63.2	1116	1	SLPH_BACBR	P38383 bacillus br	941	23	60.5	148	1	LYC2_RAT	O05820 rattus norv
869	24	63.2	1127	1	MDM1_YEAST	O58821 pyrococcus	942	23	60.5	148	1	LYC2_RAT	P73994 salmistr sci
870	24	63.2	1127	1	DNBI_EBY	O01446 saccharomyc	943	23	60.5	149	1	LYC2_RAT	O9x525 lactobacill
871	24	63.2	1127	1	MDM1_YEAST	P03327 epstein-bar	944	23	60.5	151	1	YURE_LACFE	P45371 chrometium
872	24	63.2	1128	1	POL_SFVI	P23074 simian foam	945	23	60.5	151	1	YPH7_CHRVI	O57937 methanococ
873	24	63.2	1171	1	DI43_MOUSE	O9z207 mus musculu	946	23	60.5	154	1	Y553_METJA	O48397 bacteriopho
874	24	63.2	1171	1	DI43_MOUSE	O9z207 mus musculu	947	23	60.5	154	1	GP41_BPSPI	P15845 schistosoma
875	24	63.2	1173	1	Y4CA_RHLSN	O9xkz1 pseudomonas	948	23	60.5	154	1	SM20_SCHMA	P22759 azotobacter
876	24	63.2	1197	1	CT9A_XENLA	P55383 rhizobium s	949	23	60.5	156	1	BER_AZOVI	P77914 neisseria m
877	24	63.2	1305	1	CT9A_XENLA	O14578 homo sapien	950	23	60.5	157	1	BERB_NEIGO	P56999 neisseria m
878	24	63.2	1305	1	YVCA_CAEEL	P98999 xenopus lae	951	23	60.5	157	1	UL51_HCMVA	P16792 human cytom
879	24	63.2	1327	1	CY49_CHICK	P34544 caenorhabdi	952	23	60.5	157	1	Y4UD_RHISN	P55670 rhizobium s
880	24	63.2	1334	1	RRPO_FPMY	O9d946 gallus gall	953	23	60.5	165	1	YV4E_VACCC	P20514 vaccinia vi
881	24	63.2	1335	1	RRPO_FPMY	P22168 foxtail mos	954	23	60.5	166	1	Y809_MERTA	O58219 methanococ
882	24	63.2	1353	1	CTA9_HUMAN	O60503 homo sapien	955	23	60.5	167	1	APT_BACSU	O34443 bacillus su
883	24	63.2	1353	1	CTA9_HUMAN	P51830 mus musculu	956	23	60.5	170	1	APT_BACSU	O9k4d2 bacillus ha
884	24	63.2	1402	1	BIRG_MOUSE	O9j1d3 mus musculu	957	23	60.5	171	1	APT_STAM	O32418 staphylococ
885	24	63.2	1403	1	BIRG_MOUSE	Q13075 homo sapien	958	23	60.5	172	1	PYRE_HA1N1	O9h9d2 halobacteri
886	24	63.2	1403	1	BIRG_MOUSE	O9qwk5 mus musculu	959	23	60.5	175	1	STG2_BACSU	O05409 bacillus su
887	24	63.2	1403	1	BIRG_MOUSE	O9j1d6 mus musculu	960	23	60.5	176	1	STG2_BACSU	O00861 halocarcula
888	24	63.2	1403	1	BIRG_MOUSE	P09884 homo sapien	961	23	60.5	177	1	RS13_HALMA	P41007 bacillus ca
889	24	63.2	1462	1	YOR1_YEAST	P53049 saccharomyc	962	23	60.5	179	1	PYRR_BACCL	P44772 haemophilus
890	24	63.2	1477	1	YOR1_YEAST	O04182 saccharomyc	963	23	60.5	181	1	Y459_HABIN	P39785 bacillus su
891	24	63.2	1529	1	PDRE_YEAST	Q010978 mycobacteri	964	23	60.5	181	1	Y179_THEAC	O9h1e8 thermoplas
892	24	63.2	1538	1	PSB8_MYCTU	O8xv02 ralsionia s	965	23	60.5	181	1	Y573_RICPR	Q9zcx8 rickettsia
893	24	63.2	1541	1	YH6_YEAST	P38073 saccharomyc	966	23	60.5	182	1	PAAD_ARCP2	O350934 archaeoglob
894	24	63.2	1582	1	CTRO_MOUSE	O48025 mus musculu	967	23	60.5	182	1	YPS2_SYNP2	P31556 synechococ
895	24	63.2	1597	1	YDHA_SCHPO	O93355 schizosacch	968	23	60.5	185	1	APR_CAEPL	P1455 caenorhabdi
896	24	63.2	1687	1	YDHA_SCHPO	O93355 schizosacch	969	23	60.5	185	1	YDHA_SCHPO	P38522 escherichia
897	24	63.2	1770	1	RI15_YEAST	P41621 drosophila	970	23	60.5	186	1	YB25_PYPAB	O02175 synechococ
898	24	63.2	2029	1	LAR_DROME	Q10105 schizosacch	971	23	60.5	187	1	APY_TREPA	O9xkz1 pyrococcus
899	24	63.2	2029	1	LAR_DROME	P01267 bos taurus	972	23	60.5	190	1	APY_TREPA	O84001 treponema p
900	24	63.2	2769	1	YAO5_SCHPO	O03133 saccharopol	973	23	60.5	190	1	DEF_THETH	O99w41 staphylococ
901	24	63.2	3172	1	ERY3_SACCR	O30408 b tyrocidin	974	23	60.5	192	1	Y615_PYPAB	P43522 thermus the
902	24	63.2	3587	1	TYCH_BACBR	P45444 emericella	975	23	60.5	194	1	Y615_PYPAB	O9u074 pyrococcus
903	24	63.2	4384	1	DIHA_CHLRE	O39610 chlamydomon	976	23	60.5	195	1	RTM4_MOUSE	P45761 escherichia
904	24	63.2	4499	1	ANTR_TRASC	P80110 ttrachemys s	977	23	60.5	199	1	RTM4_MOUSE	O99p72 mus musculu
905	23	60.5	52	1	ANTR_TRASC	P16652 potato viru	978	23	60.5	202	1	YF36_MTCU	O10774 mycobacteri
906	23	60.5	83	1	Y425_TREPA	O02232 halobacteri	979	23	60.5	204	1	VC05_VACCC	P34958 bacillus su
907	23	60.5	86	1	MINE_DEIRA	O83440 treponema p	980	23	60.5	204	1	VC05_VACCC	P21039 vaccinia vi
908	23	60.5	95	1	MINE_DEIRA	O9rwb8 deinococcus	981	23	60.5	204	1	VC05_VACCC	P17367 vaccinia vi
909	23	60.5	95	1	MINE_DEIRA	O9rwb8 deinococcus	982	23	60.5	204	1	VC05_VACCC	P17367 vaccinia vi

```

983      23      60.5      205      1      KGUA_CHLPN      092961 chlamydia p
984      23      60.5      205      1      KGUA_NEIMA      09196 neisseria m
985      23      60.5      205      1      KGUA_NEIMS      09195 neisseria m
986      23      60.5      207      1      ICW3_PSOTE      P10822 psophocarp
987      23      60.5      207      1      ICW3_PSOTE      09191 staphylococ
988      23      60.5      208      1      LEXA_STRAU      09191 neisseria m
989      23      60.5      208      1      LEXA_STRAU      09191 neisseria m
990      23      60.5      210      1      NORD_NEIMA      P34988 chelabact
991      23      60.5      210      1      NORD_NEIMA      P34988 chelabact
992      23      60.5      211      1      NTRA_CHEHE      P34988 chelabact
993      23      60.5      211      1      NTRA_CHEHE      P34988 chelabact
994      23      60.5      211      1      CRB3_MOUSE      P34988 chelabact
995      23      60.5      211      1      DEOC_BACSU      P34988 chelabact
996      23      60.5      212      1      PYRE_LACPL      P34988 chelabact
997      23      60.5      212      1      RAI1_RHIER      P34988 chelabact
998      23      60.5      213      1      SODE_HELPJ      P34988 chelabact
999      23      60.5      214      1      CHEZ_ECOLT      P34988 chelabact
1000     23      60.5      215      1      TDXH_THEMEA      P34988 chelabact

```

## ALIGNMENTS

```

RESULT 1
AMPA_CHLPN      STANDARD:      PRT:      499 AA.
ID      0928F8: 0920J2:
AC      30-MAY-2000 (Rel. 39, Created)
DT      15-JUN-2002 (Rel. 41, Last sequence update)
DE      Probable cytosol aminopeptidase (EC 3.4.11.1) (Leucine aminopeptidase)
DE      (LAP) (Leucyl aminopeptidase).
DE      PEPA OR CP00385 OR CP0370.
OS      Chlamydia pneumoniae (Chlamydia pneumoniae).
OC      Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX      NCBI_TaxID=83358;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      MEDLINE=99206606; PubMed=10192388;
RA      Kaiman S., Mitchell W., Marache R., Jammal C., Fan J., Hyman R.W.,
RA      Olinger L., Grilwood J., Davis R.W., Stephens R.S.;
RA      "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
RA      Nat. Genet. 21:385-389(1999).
RN      [2]
RP      SEQUENCE FROM N.A.
RC      MEDLINE=20150255; PubMed=10684935;
RA      White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
RA      Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA      Gwin M., Nelson W., Deboy R., Kolonay J., McClarty G., Salzberg S.L.,
RA      Eisen J., Fraser C.M.;
RA      "Genome sequences of Chlamydia trachomatis Mopn and Chlamydia
RA      pneumoniae AR39.";
RA      Nucleic Acids Res. 28:1397-1406(2000).
RN      [3]
RP      SEQUENCE FROM N.A.
RC      STRAIN=J138;
RA      MEDLINE=20330349; PubMed=10871362;
RA      Shirai M., Hirakawa H., Kimoto M., Tsubuchi M., Kishi F., Onchi K.,
RA      Shiba T., Ishii K., Hattori M., Kohara S., Nakazawa T.;
RA      "Comparison of whole genome sequences of Chlamydia pneumoniae J138
RA      from Japan and CWL029 from USA.";
RA      Nucleic Acids Res. 28:2311-2314(2000).
CC      -1- FUNCTION: PRESUMABLY INVOLVED IN THE PROCESSING AND REGULAR
CC      TURNOVER OF INTRACELLULAR PROTEINS. CATALYZES THE REMOVAL OF
CC      UNSUBSTITUTED N-TERMINAL AMINO ACIDS FROM VARIOUS PEPTIDES (BY
CC      SIMILARITY).
CC      -1- CATALYTIC ACTIVITY: Release of an N-terminal amino acid, Xaa-1-
CC      including Pro although not Arg or Lys, and Xbb may be Pro.
CC      -1- COFACTOR: MANGANESE (BY SIMILARITY).

```

```

CC      -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC      -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M17.
CC      -----
CC      This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@sib-sib.ch).
CC      -----
CC      EMBL: AE001623; AAD18529.1;
CC      EMBL: AE002199; AAF38219.1;
CC      EMBL: AE002546; BAA98593.1;
CC      HSSP: P00727; 1BPN.
CC      MEROPS: M17; UPM;
CC      PICT-2DPAGE: 0928F8; -.
CC      TIGR: CP0370; -.
CC      InterPro: IPR000819; Peptidase_M17.
CC      Pfam: PF00883; Peptidase_M17; 1.
CC      PRINTS: PR00481; LAMNOPTDASE.
CC      PROSITE: PS00631; CYTOSOL_AP; 1.
CC      KEGG: K01101; MANGANESE; Complete proteome.
CC      FT      METAL      263      263      MANGANESE 2 (BY SIMILARITY).
CC      FT      METAL      268      268      MANGANESE 1 AND 2 (BY SIMILARITY).
CC      FT      METAL      286      286      MANGANESE 2 (BY SIMILARITY).
CC      FT      METAL      345      345      MANGANESE 1 (BY SIMILARITY).
CC      FT      METAL      347      347      MANGANESE 1 AND 2 (BY SIMILARITY).
CC      FT      ACT SITE      275      275      POTENTIAL.
CC      FT      ACT SITE      349      349      POTENTIAL.
CC      SQ      SEQUENCE      499 AA; 54509 MW; B0F281BDF4B9EC2 CRC64;
CC      Query Match      86.8%; Score 33; DB 1; Length 499;
CC      Best Local Similarity      77.8%; Pred. NO. 21;
CC      Matches      7; Conservative      1; Mismatches      1; Indels      0; Gaps      0;
CC      Db      1 VLXDDLEA 9
CC      395 VLAEDELEA 403
CC      -----
CC      RESULT 2
CC      GINB_RHOSH      STANDARD:      PRT:      112 AA.
CC      ID      GINB_RHOSH
CC      AC      P43519;
CC      DT      01-NOV-1995 (Rel. 32, Created)
CC      DT      01-NOV-1995 (Rel. 32, Last sequence update)
CC      DT      15-DEC-1998 (Rel. 37, Last annotation update)
CC      DE      Nitrogen regulatory protein P-II (PIT signal transducing protein).
CC      GN      GINB.
CC      OS      Rhodospirillum rubrum (Rhodospirillum rubrum).
CC      OC      Bacteria; Proteobacteria; alpha subdivision; Rhodospirillum group.
CC      OX      NCBI_TaxID=1063;
CC      RN      [1]
CC      RP      SEQUENCE FROM N.A.
CC      RC      STRAIN=2R;
CC      RA      MEDLINE=95005468; PubMed=7921264;
CC      RA      Zinchenko V.V., Chutrin Y., Shestopalov V.I., Shestakov S.V.;
CC      "Nucleotide sequence and characterization of the Rhodospirillum
CC      rubrum glnB and glnA genes.";
CC      Microbiology 140:2143-2151(1994).
CC      RT      Microbiology 140:2143-2151(1994).
CC      RL      Microbiology 140:2143-2151(1994).
CC      CC      -1- FUNCTION: P-II INDIRECTLY CONTROLS THE TRANSCRIPTION OF THE
CC      GLUTAMINE SYNTHETASE GENE (GINA). P-II PREVENTS NR-II CATALYZED
CC      CONVERSION OF NR-I TO NR-I-PHOSPHATE, THE TRANSCRIPTIONAL
CC      ACTIVATOR OF GINA. WHEN P-II IS URIDYLATED TO P-II-UMP, THESE
CC      EVENTS ARE REVERSED. WHEN THE RATIO OF GIN TO 2-KETOGLUTARATE
CC      DECREASES, P-II IS URIDYLATED TO P-II-UMP, WHICH CAUSES THE
CC      DEADENYLATION OF GLUTAMINE SYNTHETASE, SO ACTIVATING THE ENZYME.
CC      -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC      -1- SIMILARITY: BELONGS TO THE P(II) PROTEIN FAMILY.

```



```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: X71659; CA50650.1; -.
CC
CC DR HSSP: P38504; 1GNK.
CC DR InterPro: IPR002187; PII_Glnb_UMP.
CC DR InterPro: IPR002187; PII_Glnb.
CC DR Pfam: PF00543; P-II; 1.
CC DR PRINTS: PR00340; PII_Glnb.
CC DR PRODOM: PD001194; PII_Glnb_UMP; 1.
CC DR PROSITE: PS00496; PII_Glnb_CTER; 1.
CC DR PROSITE: PS00638; PII_Glnb_CTER; 1.
CC Transcription regulation: Nitrogen fixation.
CC KW BINDING 51 51 UMP (BY SIMILARITY).
CC SEQUENCE 112 AA; 12100 MW; B27846AC9EB17D3 CRC64;
CC
CC Query Match 84.2%; Score 32; DB 1; Length 112;
CC Best Local Similarity 66.7%; Pred. NO. 6.8;
CC Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
CC
CC QY 1 VLXDDLEA 9
CC 11111111
CC Db 64 VLADDMVEA 72
CC
CC RESULT 3
CC CARB_METBA STANDARD; PRT; 398 AA.
CC ID CARB_METBA
CC AC P13258;
CC DT 01-JAN-1990 (Rel. 13, Created)
CC DT 01-JUN-1990 (Rel. 13, Last sequence update)
CC DT 15-JUN-2002 (Rel. 41, Last annotation update)
CC DE Carbamoyl-phosphate synthase large chain (EC 6.3.5.5) (Carbamoyl-
CC phosphate synthetase ammonia chain) (Fragment).
CC GN CARB.
CC OS Methanosarcina barkeri.
CC OC Archaea; Euryarchaeota; Methanococci; Methanosarcinales;
CC OC Methanosarcinaceae; Methanosarcina.
CC CX NCBI_TaxID=2208;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RC STRAIN=MS / DSM 800;
CC RX MEDLINE=88257029; Pubmed=3133361;
CC RA Morris C.J., Reeve J.N.;
CC RT "Conservation of structure in the human gene encoding
CC argininosuccinate lyase and the arg genes of the archaeobacteria
CC Methanosarcina barkeri MS and Methanococcus vannielii.";
CC RL J. Bacteriol. 170:3125-3130(1988).
CC CC CATALYTIC ACTIVITY: 2 ATP + L-glutamine + CO(2) + H(2)O = 2 ADP +
CC phosphate + L-glutamate + carbamoyl phosphate.
CC -I COFACTOR: Binds three manganese ions (By similarity).
CC -I PATHWAY: Arginine biosynthesis; first step.
CC -I SUBUNIT: Composed of two chains; the small (or glutamine) chain
CC promotes the hydrolysis of glutamine to ammonia, which is used by
CC the large (or ammonia) chain to synthesize carbamoyl phosphate.
CC -I SIMILARITY: BELONGS TO THE CARB FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: M21314; AAA72676.1; -.
CC PIR: A28180; A28180.

```

```

DR HSSP: P00968; 1A9X.
DR InterPro: IPR000901; CPSase.
DR InterPro: IPR004362; MGS_Like.
DR Pfam: PF02142; MGS; 1.
DR Pfam: PF02786; CPSase_L_D2; 1.
DR PROSITE: PS00866; CPSASE_1; 1.
DR PROSITE: PS00867; CPSASE_2; 1.
CC Arginine biosynthesis; Pyrimidine biosynthesis; Ligase; Manganese.
CC KW NON_TER 1 1
CC FT DOMAIN <1 253 CARBAMOYL PHOSPHATE SYNTHETIC DOMAIN.
CC FT DOMAIN 254 398 ALLOSTERIC DOMAIN.
CC FT METAL 146 146 MANGANESE 3 (BY SIMILARITY).
CC FT METAL 158 158 MANGANESE 3 (BY SIMILARITY).
CC SEQUENCE 398 AA; 43691 MW; 90C409794CAD2131 CRC64;
CC
CC Query Match 81.6%; Score 31; DB 1; Length 398;
CC Best Local Similarity 66.7%; Pred. NO. 43;
CC Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
CC
CC QY 1 VLXDDLEA 9
CC 11111111
CC Db 67 ILIDDFLEA 75
CC
CC RESULT 4
CC CARB_METMA STANDARD; PRT; 1073 AA.
CC ID CARB_METMA
CC AC P58944;
CC DT 15-JUN-2002 (Rel. 41, Created)
CC DT 15-JUN-2002 (Rel. 41, Last sequence update)
CC DT 15-JUN-2002 (Rel. 41, Last annotation update)
CC DE Carbamoyl-phosphate synthase large chain (EC 6.3.5.5) (Carbamoyl-
CC phosphate synthetase ammonia chain).
CC GN CARB OR MM0038.
CC OS Methanosarcina mazei (Methanosarcina frisia).
CC OC Archaea; Euryarchaeota; Methanococci; Methanosarcinales;
CC OC Methanosarcinaceae; Methanosarcina.
CC CX NCBI_TaxID=2209;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RC STRAIN=Goel / G01 / ATCC BAA-199 / DSM 3647 / OCM 88;
CC RX MEDLINE=22120827; Pubmed=12125824;
CC RA Deppenmeier U., Johann A., Hartsch T., Merkl R., Schmitz R.A.,
CC Martinez-Arias R., Henne A., Wietzer A., Baumeier S., Jacobi C.,
CC Brueggemann H., Lienard T., Christmann A., Boemcke M., Steckel S.,
CC Bhattacharyya A., Lykidis A., Overbeek R., Kleuk H.-P., Gunsalus R.P.,
CC Fritz H.-J., Gottschalk G.;
CC RA "The genome of Methanosarcina mazei: evidence for lateral gene
CC transfer between bacteria and Archaea.";
CC RL J. Mol. Microbiol. Biotechnol. 4:453-461(2002).
CC CC CATALYTIC ACTIVITY: 2 ATP + L-glutamine + CO(2) + H(2)O = 2 ADP +
CC phosphate + L-glutamate + carbamoyl phosphate.
CC -I COFACTOR: Binds three manganese ions (By similarity).
CC -I PATHWAY: Arginine biosynthesis.
CC -I SUBUNIT: Composed of two chains; the small (or glutamine) chain
CC promotes the hydrolysis of glutamine to ammonia, which is used by
CC the large (or ammonia) chain to synthesize carbamoyl phosphate (By
CC similarity).
CC -I SIMILARITY: BELONGS TO THE CARB FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AE013223; AAM29734.1; -.
CC DR PROSITE: PS00866; CPSASE_1; 1.
CC DR PROSITE: PS00867; CPSASE_2; 1.
CC Arginine biosynthesis; Pyrimidine biosynthesis; Ligase; Repeat;

```

KM ATP-binding; Manganese.  
 FT DOMAIN 1 399  
 FT DOMAIN 400 540  
 FT DOMAIN 541 931  
 FT DOMAIN 932 1073  
 FT REPEAT 1 540  
 FT REPEAT 541 1073  
 FT NP-BIND 153 210  
 FT NP-BIND 300 350  
 FT METAL 284 284  
 FT METAL 296 296  
 FT METAL 298 298  
 FT METAL 822 822  
 FT METAL 834 834  
 SO SEQUENCE 1073 AA: 118963 MW: ACBDE95DFCIEAD1 CRC64;

Query Match Best Local Similarity 81.6%; Score 31; DB 1; Length 1073;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 VLXDDLEA 9  
 Db 743 ILIDDFLEA 751

## RESULT 5

EPS8\_HUMAN STANDARD; PRT; 822 AA.

ID Q12929;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Epidermal growth factor receptor kinase substrate EPS8.

GN EPS8.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 NCBI\_Taxid=9606;

OX NCBI\_Taxid=9606;

RP SEQUENCE FROM N.A.

RX MEDLINE=94366758; PubMed=8084614;

RA Wong W.T., Carlomagno F., Druck T., Barletta C., Croce C.M.,

RT Huebner K., Kraus M.H., di Fiore P.P.;

RT "Evolutionary conservation of the EPS8 gene and its mapping to human

chromosome 12q23-q24.";

RT Oncogene 9:3057-3061(1994).

CC - FUNCTION: UPON BINDING TO EGF RECEPTOR ENHANCES EGF-DEPENDENT

MITOGENIC SIGNALS. CAN BIND MULTIPLE CELLULAR TARGETS.

CC - TISSUE SPECIFICITY: EXPRESSED IN ALL TISSUES ANALYZED, INCLUDING

HEART, BRAIN, PLACENTA, LUNG, LIVER, SKELETAL MUSCLE, KIDNEY AND

PANCREAS. EXPRESSED IN ALL EPITHELIAL AND FIBROBLASTIC LINES

CC EXAMINED AND IN SOME, BUT NOT ALL, HEMATOPOIETIC CELLS.

CC - PTM: PHOSPHORYLATED BY SEVERAL RECEPTOR TYROSINE KINASES.

CC - SIMILARITY: CONTAINS 1 SH3 DOMAIN.

CC - SIMILARITY: CONTAINS 1 PH DOMAIN. IT IS SPLIT IN TWO PARTS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation

at the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: U1535; AAA62280.1; -

DR HSSP: 008509; 1AOT

DR Genew: HGNC:3420; EPS8.

DR MW: 600206; -

DR InterPro: IPR000050; PID\_domain.

DR InterPro: IPR001452; SH3.

DR Pfam: PF00018; SH3; 1.

DR ProDom: PD000066; SH3; 1.

DR SMART: SM00462; PTB; 1.

DR SMART: SM00326; SH3; 1.  
 DR PROSITE: PS50002; SH3; 1.  
 KM SH3 domain; Phosphorylation.  
 FT DOMAIN 69 129  
 FT DOMAIN 129 213  
 FT DOMAIN 213 325  
 FT DOMAIN 325 414  
 FT DOMAIN 414 440  
 FT DOMAIN 440 532  
 FT DOMAIN 532 591  
 FT DOMAIN 591 651  
 FT DOMAIN 651 664  
 SO SEQUENCE 822 AA: 91881 MW: ACSEBID28B784B3B CRC64;

Query Match Best Local Similarity 78.9%; Score 30; DB 1; Length 822;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLXDDLE 8  
 Db 552 VLKDDLE 559

## RESULT 6

YBDA\_ECOLI

ID YBDA\_ECOLI

AC P15050;

DT 01-APR-1990 (Rel. 14, Created)

DT 01-APR-1990 (Rel. 14, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Hypothetical protein ybda (p15).

GN YBDA OR B0597

OS Escherichia coli.

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

OC Escherichia.

OX NCBI\_Taxid=562;

RP SEQUENCE FROM N.A.

RX MEDLINE=89123155; PubMed=2521622;

RA Liu J., Duncan K., Walsh C.T.;

RT "Nucleotide sequence of a cluster of Escherichia coli enterobactin

biosynthesis genes: identification of enta and purification of its

product 2,3-dihydro-2,3-dihydroxybenzoate dehydrogenase.";

RT J. Bacteriol. 171:791-798(1989).

CC - FUNCTION: UPON BINDING TO EGF RECEPTOR ENHANCES EGF-DEPENDENT

MITOGENIC SIGNALS. CAN BIND MULTIPLE CELLULAR TARGETS.

CC - TISSUE SPECIFICITY: EXPRESSED IN ALL TISSUES ANALYZED, INCLUDING

HEART, BRAIN, PLACENTA, LUNG, LIVER, SKELETAL MUSCLE, KIDNEY AND

PANCREAS. EXPRESSED IN ALL EPITHELIAL AND FIBROBLASTIC LINES

CC EXAMINED AND IN SOME, BUT NOT ALL, HEMATOPOIETIC CELLS.

CC - PTM: PHOSPHORYLATED BY SEVERAL RECEPTOR TYROSINE KINASES.

CC - SIMILARITY: CONTAINS 1 SH3 DOMAIN.

CC - SIMILARITY: CONTAINS 1 PH DOMAIN. IT IS SPLIT IN TWO PARTS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation

at the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: U1535; AAA62280.1; -

DR HSSP: 008509; 1AOT

DR Genew: HGNC:3420; EPS8.

DR MW: 600206; -

DR InterPro: IPR000050; PID\_domain.

DR InterPro: IPR001452; SH3.

DR Pfam: PF00018; SH3; 1.

DR ProDom: PD000066; SH3; 1.

DR SMART: SM00462; PTB; 1.



entitles requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL: M24148; AAA16104.1; -  
 EMBL: M24143; AAA76837.1; -  
 EMBL: AE000165; AAC73698.1; -  
 EMBL: U82598; AAB40797.1; -  
 PIR: D32047; Q3CECA.  
 Ecogene; EG1105; ybdb.  
 InterPro; IPR003736; DUF157.  
 Pfam; PF02584; DUF157; 1.  
 TIGRPFAMs; TIGR00369; unchar.dom.1; 1.  
 Hypothetical protein; Complete proteome.  
 SEQUENCE 137 AA; 14970 MW; C8DF8DE63815F206 CRC64;

Query Match  
 Best Local Similarity 76.3%; Score 29; DB 1; Length 137;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LXDLEA 9  
 111111  
 Db 31 LGDDVLEA 38

RESULT 7  
 RL30\_HALMA STANDARD; PRT; 154 AA.  
 AC P14121;  
 DT 01-JAN-1990 (Rel. 13, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE 50S ribosomal protein L30P (Hma130) (HL20) (HL16).  
 GN RPL30P.  
 OS Halococcus marismortui (Halobacterium marismortui).  
 OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;  
 OC Halobacteriaceae; Halocarcula.  
 NCBI\_TaxID=2238;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91360093; PubMed=1832208;  
 RA Scholzen T., Arndt E.;  
 RT "Organization and nucleotide sequence of ten ribosomal protein genes from the region equivalent to the spectinomycin operon in the archaeobacterium Halobacterium marismortui.";  
 RT Mol. Gen. Genet. 228:70-80(1991).  
 RL [2]  
 RP SEQUENCE.  
 RX MEDLINE=30076190; PubMed=2591382;  
 RA Hatakeyama T., Kaufmann F., Schroeter B., Hatakeyama T.;  
 RT "Primary structures of five ribosomal proteins from the archaeobacterium Halobacterium marismortui and their structural relationships to eubacterial and eukaryotic ribosomal proteins.";  
 RT Eur. J. Biochem. 185:685-693(1989).  
 RL [3]  
 RP SEQUENCE OF 1-30.  
 RX MEDLINE=89062418; PubMed=3196689;  
 RA Walsh M.J., McDougall J., Wiltmann-Liebold B.;  
 RT "Extended N-terminal sequencing of archaeobacterial ribosomes blotted from two-dimensional gels onto glass fiber and poly(vinylidene difluoride) membrane.";  
 RT Biochemistry 27:6867-6876(1988).  
 RL [4]  
 RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS).  
 RC STRAIN=ATCC 43049;  
 RX MEDLINE=20396344; PubMed=10937989;  
 RA Ban N., Nissen P., Hansen J., Moore P.B., Steitz T.A.;  
 RT "The complete atomic structure of the large ribosomal subunit at 2.4 A resolution.";  
 RT Science 289:905-920(2000).  
 RL -1- SIMILARITY: BELONGS TO THE L30P FAMILY OF RIBOSOMAL PROTEINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL: X58395; CAA41292.1; -  
 PIR: S16543; R5HS30.  
 PDB; 1FFK; 14-ANG-00.  
 InterPro; IPR000517; Ribosomal\_L30.  
 Pfam; PF00327; Ribosomal\_L30; 1.  
 TIGRPFAMs; TIGR01309; L30P\_arch; 1.  
 PROSITE; PS00634; RIBOSOMAL\_L30; 1.  
 Ribosomal protein; 3D-structure.  
 CONFLICT 83 W -> L (IN REF. 2).  
 FT CONFLICT 148 148 MISSING (IN REF. 2).  
 FT SEQUENCE 154 AA; 17042 MW; ED673F036E974C14 CRC64;

Query Match  
 Best Local Similarity 76.3%; Score 29; DB 1; Length 154;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 DDLEA 9  
 111111  
 Db 147 DDLEA 152

RESULT 8  
 YHWM\_ECO57 STANDARD; PRT; 231 AA.  
 ID YHWM\_ECO57  
 AC P58116;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 15-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Protein yhmW.  
 GN YHWM OR Z4807 OR ECS4288.  
 OS Escherichia coli O157:H7.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.  
 NCBI\_TaxID=83334;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=O157:H7 / EDU933 / ATCC 700927;  
 RX MEDLINE=21074935; PubMed=11206551;  
 RA Perna N.T., Plunkett G., III, Burland V., Mau B., Glasner J.D., Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A., Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L., Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K., Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C., Welch R.A., Blattner F.R.;  
 RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";  
 RT Nature 409:529-533(2001).  
 RL [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=O157:H7 / RUMD 0509952;  
 RX MEDLINE=21156231; PubMed=11258796;  
 RA Hayashi T., Makino K., Onishi M., Kurokawa K., Ishii K., Yokoyama K., Han C.-G., Ohtsuka E., Nakayama K., Murata T., Tanaka M., Tobe T., Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T., Kohara S., Shiba T., Hattori M., Shinagawa H.;  
 RT "Complete genome sequence of enterohaemorrhagic Escherichia coli O157:H7 and genomic comparison with a laboratory strain K-12.";  
 RT DNA Res. 8:11-22(2001).  
 RL -1- SIMILARITY: BELONGS TO THE PIRIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: AE005567; AAG58548.1; -  
 DR EMBL: AP002565; BAB37711.1; -  
 DR InterPro: IPR003829; DUF209.  
 DR Pfam: PF02678; DUF209; 1.  
 KW Complete proteome.

SEQUENCE 231 AA; 26291 MW; DC15B213316875E4 CRC64;

Query Match 76.3%; Score 29; DB 1; Length 231;  
 Best Local Similarity 55.6%; Pred. No. 61;

Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 VLXDDLLEA 9  
 I: ||: ||  
 DB 42 VINDVIEA 50

## RESULT 9

YHWM\_ECOLI STANDARD; PRT; 231 AA.

AC P46852.

DT 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Protein yhmw.

GN YHWM OR B3439.

OS Escherichia coli.

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

OC Escherichia.

OX NCBI\_Taxid=562;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-K12 / MG1655;

RX MEDLINE=9742617; PubMed=9278503;

RA Blatter F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,

RA Riley M., Collado-VIDES J., Glasner J.D., Rode C.K., Mayhew G.F.,

RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,

RA Mau B., Shao Y.

RT "The complete genome sequence of Escherichia coli K-12."

RT Science 277:1453-1474(1997).

RL [2]

RP IDENTIFICATION BY MASS SPECTROMETRY.

RX MEDLINE=99420866; PubMed=10493123;

RA Pountoulakis M., Takacs M.-F., Berndt P., Langen H., Takacs B.,

RT "Enrichment of low abundance proteins of Escherichia coli by

RT hydroxyapatite chromatography."

RT Electrophoresis 20:2181-2195(1999).

RL [1]

CC - SIMILARITY: BELONGS TO THE PIRIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

RESULT 10  
 COX2\_CANCA STANDARD; PRT; 251 AA.  
 ID COX2\_CANCA  
 AC P43733;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Cytochrome c oxidase polypeptide II (EC 1.9.3.1).  
 GN COX2.  
 OS Candida glabrata (Yeast) (Torulopsis glabrata).  
 OG Mitochondrion.  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; mitosporic Saccharomycetales; candida.  
 OX NCBI\_Taxid=5478;

SEQUENCE FROM N.A.

RP STRAIN-CBS 138;

RX MEDLINE=94365843; PubMed=8083884;

RA Clark-Walker G.D., Meller G.F.;

RT "The structure of the small mitochondrial DNA of Kluyveromyces

thermotolerans is likely to reflect the ancestral gene order in

fungi."

RT J. Mol. Evol. 38:593-601(1994).

RL [1]

CC - FUNCTION: CYTOCHROME C OXIDASE IS THE COMPONENT OF THE RESPIRATORY

CHAIN THAT CATALYZES THE REDUCTION OF OXYGEN TO WATER. SUBUNIT 1-

3 FORM THE FUNCTIONAL CORE OF THE ENZYME COMPLEX. SUBUNIT 2

TRANSFERS THE ELECTRONS FROM CYTOCHROME C VIA ITS BINDING COPPER

A CENTER TO THE BIMETALLIC CENTER OF THE CATALYTIC SUBUNIT 1.

CC - CATALYTIC ACTIVITY: 4 ferrioxochrome c + O(2) = 4 ferrioxochrome

c + 2 H(2)O.

CC - COFACTOR: Copper A.

CC - SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial

inner membrane.

CC - SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE SUBUNIT 2 FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: X69430; CAA9205.1; -

DR InterPro: IPR001505; Copper\_CuA.

DR InterPro: IPR002429; Cyt\_cox\_2.

DR Pfam: PF00116; COX2; 1.

DR Pfam: PF02790; COX2\_TM; 1.

DR PRINTS: PR01166; CYCOXIDASEII.

DR PRODOM: PD000131; Copper\_CuA; 1.

DR PROSITE: PS00078; COX2; 1.

KW Oxidoreductase; Copper; Mitochondrion; Transmembrane; Inner membrane;

KW Electron transport; Respiratory chain.

FT METAL 186 186 COPPER A (PROBABLE).

FT METAL 221 221 COPPER A (PROBABLE).

FT METAL 225 225 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229 COPPER A (PROBABLE).

FT METAL 229 229

```

DR 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Cytochrome c1, heme protein, mitochondrial precursor.
GN CYT1.
OS Kluveromyces lactis (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Kluveromyces.
OX NCBI_TaxID=28985;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JBD100;
RX MEDLINE=96304292; PubMed=8660461;
RA Gbeleska Y., Horvathova K., van der Aart Q.J.M., Zonneveld B.,
RA Steensma H.Y., Subik J.;
RT "Isolation and molecular analysis of the gene for cytochrome c1 from
RT Kluveromyces lactis."
RT Curr. Genet. 30:145-150(1996).
CC - FUNCTION: THIS IS THE HEME-CONTAINING COMPONENT OF THE CYTOCHROME
CC B-C1 COMPLEX, WHICH ACCEPTS ELECTRONS FROM RIESKE PROTEIN AND
CC TRANSFERS ELECTRONS TO CYTOCHROME C IN THE MITOCHONDRIAL
CC RESPIRATORY CHAIN.
CC - SUBUNIT: THE MAIN SUBUNITS OF COMPLEX B-C1 ARE: CYTOCHROME B,
CC CYTOCHROME C1 AND THE RIESKE PROTEIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X95899; CAA65144.1;
DR InterPro: IPR000345; Cyto_c1_heme_bind.
DR InterPro: IPR002326; Cyto_c1.
DR Pfam: PF02167; Cytochrome_c1.1.
DR PRINTS: PR00603; CYTOCHROME_C1.
DR PROSITE: PS00190; CYTOCHROME_C; 1.
DR Electron transport, Respiratory chain; Oxidative phosphorylation;
DR Heme; Mitochondrion; Transit peptide.
RW TRANSIT 1 46 MITOCHONDRION (POTENTIAL).
FT CHAIN 47 292 CYTOCHROME C1, HEME PROTEIN.
FT BINDING 86 86 HEME (COVALENT).
FT BINDING 89 89 HEME (COVALENT).
FT METAL 90 90 IRON (HEME AXIAL LIGAND).
FT METAL 210 210 IRON (HEME AXIAL LIGAND). (BY SIMILARITY).
SQ SEQUENCE 292 AA; 31947 MW; 07F909B81B1C1C72 CRC64;

Query Match 76.3%; Score 29; DB 1; Length 292;
Best Local Similarity 75.0%; Pred. No. 79;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 LKDDLE 8
DB 213 VLFDLVE 220

RESULT 12
YMX7_CAEEL STANDARD; PRT; 322 AA.
AC P34515;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical 37.6 kDa protein K06H7.7 in chromosome III.
GN K06H7.7.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;

```

```

RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Alnsough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Pavello A., Fraser A.,
RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
RA Johnston L., Jones M., Kershaw J., Kirsten J., Laister N.,
RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,
RA Sims M., Smailson N., Smith A., Smith M., Sonhammer E., Staden R.,
RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
RA Waterston R., Watson A., Weinstein L., Wilkinson-Sproat J.,
RA Wohldman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans."
RT Nature 368:32-38(1994).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L15314; AAF99983.1;
DR PIR: S44847; S44847.
DR WormPep: K06H7.7; CE00258.
KW Hypothetical protein.
SQ SEQUENCE 322 AA; 37625 MW; D3A874DCEFC0176D CRC64;

Query Match 76.3%; Score 29; DB 1; Length 322;
Best Local Similarity 85.7%; Pred. No. 88;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 LKDDLE 8
DB 221 LADLLE 227

RESULT 13
MSBB_ECOLI STANDARD; PRT; 323 AA.
ID MSBB_ECOLI
AC P24205;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Lipid A biosynthesis (KDO)2-(lauroyl)-lipid IVA acyltransferase
DE (EC 2.3.1.-).
GN MSBB OR B1855.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / W3110;
RA MEDLINE=92121107; PubMed=1732206;
RA Kato M., Georgopoulos C.;
RT "Isolation and characterization of the Escherichia coli msbB gene, a
RT multicopy suppressor of null mutations in the high-temperature
RT requirement gene htrB."
RT J. Bacteriol. 174:702-710(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=93015688; PubMed=1356966;
RA Engel H., Smink A.J., van Wijngaarden L., Keck W.;
RT "Murein-metabolizing enzymes from Escherichia coli: existence of a
RT second lytic transglycosylase."
RT J. Bacteriol. 174:6394-6403(1992).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;

```

RX MEDLINE=97426617; PubMed=9278503;  
 RA Blatner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
 RA Mau B., Shao Y.;  
 RT "The complete genome sequence of *Escherichia coli* K-12.";  
 RN Science 277:1453-1474(1997).  
 [4]  
 RC SEQUENCE FROM N.A.  
 RP STRAIN=K12;  
 RX MEDLINE=97251358; PubMed=9097040;  
 RA Itoh T., Alba H., Baba T., Fujita K., Hayashi K., Inada T.,  
 RA Isono K., Kasai H., Kimura S., Kitakawa M., Kitagawa M.,  
 RA Makino K., Miki T., Mizobuchi K., Mori H., Mori T., Motomura K.,  
 RA Nakade S., Nakamura Y., Nishimoto H., Nishio Y., Osihima T.,  
 RA Saito N., Sempel G., Seki Y., Sivasubramanian S., Tagami H.,  
 RA Takeda J., Takemoto K., Wada C., Yamamoto Y., Horikuchi T.;  
 RT "A 460-kb DNA sequence of the *Escherichia coli* K-12 genome  
 corresponding to the 40.1-50.0 min region on the linkage map.";  
 RN DNA Res. 3:379-392(1996).  
 [5]  
 RP FUNCTION, AND CHARACTERIZATION.  
 RX MEDLINE=97256743; PubMed=9099672;  
 RA Clementz T., Zhou Z., Raetz C.R.H.;  
 RT "Function of the *Escherichia coli* msbA gene, a multicomponent suppressor  
 of htrb knockouts, in the acylation of lipid A. Acylation by Mabb  
 follows laurate incorporation by Htrb.";  
 RL J. Biol. Chem. 272:10353-10360(1997).  
 CC -1- FUNCTION: TRANSFERS MYRISTATE OR LAURATE, ACTIVATED ON ACP, TO THE  
 CC LIPID IVA MOIETY OF (KDO)2-(LAUROYL)-LIPID IVA. DECAOYL,  
 CC PALMITOYL, PALMITOLEYL, AND (R)-3-HYDROXYMYRISTOYL-ACP ARE POOR  
 CC ACYL DONORS. FUNCTIONS OPTIMALLY AFTER LAURATE INCORPORATION BY  
 CC HTRB HAS TAKEN PLACE. ACYLATES (KDO)2-(LAUROYL)-LIPID IVA ABOUT  
 CC 100 TIMES FASTER THAN (KDO)2-LIPID IVA. DISPLAYS A PREFERENCE FOR  
 CC MYRISTOYL-ACP OVER LAUROYL-ACP.  
 CC -1- PATHWAY: Lipopolysaccharide core biosynthesis.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.  
 CC -1- SIMILARITY: BELONGS TO THE HTRB/MSB FAMILY.  
 CC -1- CAUTION: WAS ORIGINALLY (REF.2) THOUGHT TO BE THE MEMBRANE-BOUND  
 CC LYTIC MORELIN TRANSGLYCOSYLASE (MLT).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL: M77039; AAA24181.1;  
 DR EMBL: M87660; AAA96706.1;  
 DR EMBL: AE000279; AAC74925.1;  
 DR EMBL: D90828; BAA15663.1;  
 DR PIR: A42608; A42608.  
 DR Ecocore: EG10614; msbA.  
 DR InterPro: IPR004960; LipA\_acyltans.  
 DR Pfam: PF03279; LipA\_acyltans; 1.  
 KW Lipopolysaccharide biosynthesis; Transferase; Acyltransferase;  
 KW Transmembrane; Inner membrane; Complete proteome.  
 FT TRANSMEM 23 43 POTENTIAL.  
 FT TRANSMEM 85 105 POTENTIAL.  
 FT TRANSMEM 133 153 POTENTIAL.  
 SQ SEQUENCE 323 AA; 37410 MW; 94DNF38A757D20CD CRC64;  
 Query Match 76.3%; Score 29; DB 1; Length 323;  
 Best Local Similarity 100.0%; Pred. No. 88;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 14  
 VIND\_BACSU STANDARD; PRT; 391 AA.  
 AC P80861; Q9R442;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE NADH dehydrogenase-like protein yjID (EC 1.6.99.-) (Glucose  
 DE starvation-inducible protein 5) (GST5).  
 GN yjID.  
 OS Bacillus subtilis.  
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
 OX NCBI\_TaxID=1423;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168;  
 RX MEDLINE=98240225; PubMed=9579062;  
 RA Rivolta C., Soldo B., Lazarevic V., Joris B., Mauel C., Karamata D.;  
 RT "A 35.7 kb DNA fragment from the *Bacillus subtilis* chromosome  
 containing a putative 12.3 kb operon involved in hexuronate catabolism  
 and a perfectly symmetrical hypothetical catabolite-responsive  
 element.";  
 RL Microbiology 144:877-884(1998).  
 [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168;  
 RX MEDLINE=98044033; PubMed=9384377;  
 RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,  
 RA Azevedo V., Berrtero M.G., Bessieres P., Bolochin A., Borchert S.,  
 RA Boriss R., Bourrier L., Brans A., Braun M., Brignell S.C., Bron S.,  
 RA Brocillet S., Brusch C.V., Caldwell B., Capuano V., Carter N.M.,  
 RA Choi S.K., Codani J.J., Conerton I.F., Cummings S.D., Daniel R.A.,  
 RA Deniro F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,  
 RA Britton K.D., Erttington J., Fabret C., Ferrari E., Foulger D.,  
 RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,  
 RA Chim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,  
 RA Gilmispi G., Guy B.J., Haga K., Haiech J., Hatwood C.R., Henaut A.,  
 RA Hilbert H., Holtsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,  
 RA Joris B., Karamata D., Kasahara Y., Kleier-Blanchard M., Klein C.,  
 RA Kobayashi Y., Koeltter P., Koningsstein G., Krogh S., Kunano M.,  
 RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,  
 RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,  
 RA Medina N., Mellado R.P., Mizuno M., Mostl D., Nakai S., Noback M.,  
 RA Moore D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,  
 RA Parro V., Pohl T.M., Portecelle D., Porwollik S., Prescott A.M.,  
 RA Presecan E., Puig P., Purnelle B., Rapoport G., Rey M., Reynolds S.,  
 RA Rieger M., Rivolta C., Roeha E., Roche B., Rose M., Sadate Y.,  
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,  
 RA Sekiguchi J., Sekowska A., Seror S.J., Serron P., Shin B.S., Soldo B.,  
 RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,  
 RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,  
 RA Tosio V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,  
 RA Viari A., Wambut R., Wedler E., Wedler H., Weitzenecker T.,  
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasunoto K., Yata K.,  
 RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.;  
 RT "The complete genome sequence of the Gram positive bacterium *Bacillus*  
 RT subtilis.";  
 RL Nature 390:249-256(1997).  
 [3]  
 RP SEQUENCE OF 1-10.  
 RC STRAIN=168 / IS58;  
 RX MEDLINE=97443988; PubMed=9298659;  
 RA Antelmann H., Bernhardt J., Schmid R., Mach H., Voelker U.,  
 RA Hecker M.;  
 RT "First steps from a two-dimensional protein index towards a response-  
 RT regulation map for *Bacillus subtilis*.";  
 RL Electrophoresis 18:1451-1463(1997).  
 CC -1- COFACTOR: FAD (BY SIMILARITY).  
 CC -1- INDUCTION: BY GLUCOSE STARVATION.  
 CC -1- SIMILARITY: BELONGS TO THE NADH DEHYDROGENASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -

CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 CC EMBL: AF015825; AAC64325.1; -  
 CC DR EMBL: 299110; CAB13086.1; -  
 CC DR Subtilist: BG13203; yJ1D.  
 CC DR InterPro: IPR001327; FAD\_pyr\_redox.  
 CC DR InterPro: IPR001100; Pyr\_redox.  
 CC DR Pfam: PF00070; pyr\_redox.1.  
 CC DR PRINTS: PR00366; FADPNR.  
 CC DR PRINTS: PR00411; PNDPRTASE1.  
 CC DR PRODOM: PD000139; FAD\_pyr\_redox.1.  
 CC KW Oxidoreductase; Flavo-protein; FAD; Complete proteome.  
 CC FT INT MET 0  
 CC SEQUENCE 391 AA; 41822 MW; E89FE873CB25455B CRC64;

Query Match 76.3%; Score 29; DB 1; Length 391;  
 Best Local Similarity 75.0%; Pred. No. 1.1e+02;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 VLXDDLE 8  
 Db 204 VLPPDLIT 211

RESULT 15  
 VATH\_DROME STANDARD; PRT; 392 AA.  
 AC 09VJ1; 09VJ2;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 15-OCT-2001 (Rel. 41, Last sequence update)  
 DE Vacuolar ATP synthase subunit H (EC 3.6.3.14) (V-ATPase H subunit)  
 DE (Vacuolar proton pump H subunit) (Vacuolar proton pump subunit SFD).  
 GN VHA5FD OR CG17996 OR CG17332.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;  
 OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;  
 OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
 OC NCBL\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Westhoff M.A., Dow J.A.T.;  
 RT "Characterization of vha5fd, the gene encoding a SFD subunit of the  
 RL Drosophila V-ATPase";  
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.  
 RL [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Berkeley;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.H.C., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abell J.F., Agbayani A., An H.-J., Andrews-Frankoch C., Baldwin D.,  
 RA Baller R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Butler H., Cadieu L.B., Davies P.,  
 RA Butulis K.C., Busam D.A., Butler H., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Foster C., Gabriellian A.E., Gary N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris K.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howard T.J., Wei M.-H., Ibegwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,

RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Morklov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon R., Nusskern D.R., Pacle J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stappleton M., Strong R., Sun E.,  
 RA Svirskaas R., Tector C., Turner R., Ventler E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissensbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RA "The genome sequence of Drosophila melanogaster.";  
 RL Science 287:2185-2195(2000).  
 CC -1- FUNCTION: SUBUNIT OF THE PERIPHERAL V1 COMPLEX OF VACUOLAR ATPASE.  
 CC SUBUNIT H ACTIVATES ATPASE ACTIVITY OF THE ENZYME AND COUPLE ATPASE  
 CC ACTIVITY TO PROTON FLOW. VACUOLAR ATPASE IS RESPONSIBLE FOR  
 CC ACIDIFYING A VARIETY OF INTRACELLULAR COMPARTMENTS IN EUKARYOTIC  
 CC CELLS, THUS PROVIDING MOST OF THE ENERGY REQUIRED FOR TRANSPORT  
 CC PROCESSES IN THE VACUOLAR SYSTEM (BY SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (in) = ADP + phosphate +  
 CC H(+) (out).  
 CC -1- SUBUNIT: V-ATPase is a heteromultimeric enzyme composed of a  
 CC peripheral catalytic V1 complex (components A to H) attached to a  
 CC an integral membrane V0 proton pore complex (components: a, c, c',  
 CC c", and d).  
 CC -1- SIMILARITY: BELONGS TO THE V-ATPASE H SUBUNIT FAMILY.  
 CC -1- CAUTION: Ref.2 sequence differs from that shown due to erroneous  
 CC gene model prediction.

CC -----  
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 CC EMBL: AF159457; AAD47254.1; -  
 CC DR EMBL: AE003652; AAF53555.2; ALT\_SEQ.  
 CC DR EMBL: AE003652; AAF53555.2; ALT\_SEQ.  
 CC DR FLYBASE: FBgn0027779; Vha5FD.  
 CC DR InterPro: IPR004908; V-ATPase\_H.  
 CC DR Pfam: PF03224; V-ATPase\_H.1.  
 CC KW Hydrolyase; ATP synthetase; Hydrogen ion transport.  
 CC SEQUENCE 392 AA; 45159 MW; B7FE1BB712F4305 CRC64;

Query Match 76.3%; Score 29; DB 1; Length 392;  
 Best Local Similarity 75.0%; Pred. No. 1.1e+02;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 VLXDDLE 8  
 Db 79 VLPPDLIT 86

RESULT 16  
 Y284\_AQUAE STANDARD; PRT; 440 AA.  
 ID Y284\_AQUAE  
 AC 066638;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Hypothetical protein AQ\_284.  
 GN AQ\_284.  
 OS Aquifex aeolicus.  
 CC Bacteria; Aquificae; Aquificae (class); Aquificales; Aquificaceae;  
 CC Aquifex.  
 CC NCBL\_TaxID=63363;

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-VF5;
RX MEDLINE=98196666; PubMed-9537320;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.R., Overbeek R., Sneed M.A., Keller M., Anjaj M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
RT aeolicus."
RL Nature 392:353-358(1998).
CC -1- SIMILARITY: BELONGS TO THE UPF0004 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch.)
CC
DR EMBL: AF000682; AAC06605.1; -
DR InterPro: IPR002792; TRAM.
DR InterPro: IPR001861; UPF0004.
DR Pfam: PF00919; UPF0004; 1.
DR Pfam: PF01938; TRAM; 1.
DR TIGRfam: TIGR00089; UPF0004; 1.
DR PROSITE: PS01278; UPF0004; 1.
DR Hypothetical protein, complete proteome.
KW SEQUENCE 440 AA; 50321 MW; 0F945FB1D7BA330 CRC64;

Query Match
Best Local Similarity 76.3%; Score 29; DB 1; Length 440;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 LKXDLLEA 9
DB 243 LKXDLLEA 250

RESULT 17
ID YS02_CAEEL STANDARD; PRT; 495 AA.
AC 009357;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 55.5 kDa protein ZK1128.2 in chromosome III.
GN ZK1128.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Bristol N2;
RA Berks M.;
RL Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE UPF0049 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch.)
CC
DR EMBL: Z47357; CA87421.1; -
DR WormRep: ZK1128.2; CE01684.
KW Hypothetical protein.
SQ SEQUENCE 495 AA; 55549 MW; BCB23D276618804E CRC64;

Query Match
Best Local Similarity 76.3%; Score 29; DB 1; Length 495;

```

```

Best Local Similarity 66.7%; Pred. No. 1.4e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLXDLLEA 9
DB 87 LKXDLLEA 95

RESULT 18
ID AMYB_TRIRP STANDARD; PRT; 496 AA.
AC 065015;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DE Beta-amylose (EC 3.2.1.2) (1,4-alpha-D-glucan maltohydrolase).
GN BMY1.
OS Trifolium repens (Creeping white clover).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicot; Rosidae;
OC Eurosid 1; Fabales; Fabaceae; Papilionoideae; Trifolieae; Trifolium.
RN [1]
RP SEQUENCE FROM N.A.
RA Gallagher J., Gana J.A., Pollock C., Cunningham S.M., Volenec J.J.;
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: CATALYZES THE LIBERATION OF MALTOSE FROM 1,4-ALPHA-D-
CC GLUCANS.
CC -1- CATALYTIC ACTIVITY: Hydrolysis of 1,4-alpha-glucosidic linkages in
CC polysaccharides so as to remove successive maltose units from the
CC non-reducing ends of the chains.
CC -1- SIMILARITY: BELONGS TO FAMILY 14 OF GLYCOSYL HYDROLASES
CC (BETA-AMYLASES).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch.)
CC
DR EMBL: AF049098; AAD04259.1; -
DR HSP: P10538; IBTC.
DR InterPro: IPR001554; GH_14.
DR Pfam: PF01373; Glyco_hydro_14; 1.
DR PRINTS: PR00750; BETAAMYLASE.
DR PROSITE: PS00506; BETA_AMYLASE_1; 1.
DR PROSITE: PS00679; BETA_AMYLASE_2; 1.
KW Hydrolyase; Glycosidase; Polysaccharide degradation.
FT ACT_SITE 102 102
FT ACT_SITE 187 187
FT ACT_SITE 187 187
SQ SEQUENCE 496 AA; 56088 MW; 8FC446AB7C927F19 CRC64;

Query Match
Best Local Similarity 76.3%; Score 29; DB 1; Length 496;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 DDLLEA 9
DB 472 DDLLEA 477

RESULT 19
ID LEPA_BORBU STANDARD; PRT; 606 AA.
AC 051115;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE GTP-binding protein lepa.
GN LEPA OR B80088.
OS Borrelia burgdorferi (Lyme disease spirochete).

```

```

OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX NCBI_TaxID=139;
RN [1]
RC SEQUENCE FROM N.A.
RA STRAIN=ATCC 35210 / B31;
RX MEDLINE=98065943; PubMed=9403685;
RA Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
RA Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
RA Peterson J., Kierlavage A.R., Quackenbush J., Salzberg S., Hanson M.,
RA van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
RA Uitterback T., Matthey L., McDonald L., Artlich P., Bowman C.,
RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
RA Smith H.O., Venter J.C.;
RT "genomic sequence of a Lyme disease spirochaete, Borrelia
RT burgdorferi";
RL Nature 390:580-586(1997).
CC -1 SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
CC LEPA SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF001121; AAC66469.1; ALT_INIT.
DR HSSP; P13551; ILEO.
DR TIGR; BB0088;
DR InterPro: IPR004161; EFTU_D2.
DR InterPro: IPR000795; EF_GTPbind.
DR InterPro: IPR005225; Small_GTP.
DR Pfam: PF00009; GTP_EFTU; 1.
DR Pfam: PF03144; GTP_EFTU_D2; 1.
DR TIGRfam: TIGR00231; small_GTP; 1.
DR PROSITE; PS00301; EFATOR_GTP; 1.
DR KW GTP-binding; Complete proteome.
FT NP_BIND 19 26 GTP (BY SIMILARITY).
FT NP_BIND 85 89 GTP (BY SIMILARITY).
FT NP_BIND 139 142 GTP (BY SIMILARITY).
SQ SEQUENCE 606 AA; 68172 MW; 29E9F10445246255 CRC64;

Query Match
Best Local Similarity 76.3%; Score 29; DB 1; Length 606;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 DDLLEA 9
DB 178 DDLLEA 183

RESULT 20
CPSA_MOUSE STANDARD; PRT; 1441 AA.
AC 09EP04;
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Cleavage and polyadenylation specificity factor, 160 kDa subunit (CPSP
DE 160 kDa subunit).
GN CPSF1 OR CPSP160.
OS Mus musculus (mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
ON NCBI_TaxID=10090;
RX MEDLINE=9403685; PubMed=9403685;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Das B., Attaya E.N., Michelle Wallace A., Macdonald C.C.;
RT "Overexpression of the Cstf-64 and CPSP-160 polyadenylation protein

```

```

RT messenger RNAs in mouse male germ cells."
RL Biol. Reprod. 64:1722-1729(2001).
CC -1 FUNCTION: CPSP PLAYS A KEY ROLE IN PRE-MRNA 3'-END FORMATION,
CC RECOGNIZING THE AAUAAA SIGNAL, SEQUENCE AND INTERACTING WITH
CC POLY(A) POLYMERASE AND OTHER FACTORS TO BRING ABOUT CLEAVAGE AND
CC POLY(A) ADDITION. THIS SUBUNIT IS INVOLVED IN THE RNA RECOGNITION
CC STEP OF THE POLYADENYLATION REACTION (By similarity).
CC -1 SUBUNIT: CPSP IS A HETEROTETRAMER COMPOSED OF FOUR DISTINCT
CC SUBUNITS 160, 100, 70 AND 30 kDa.
CC -1 SUBCELLULAR LOCATION: NUCLEAR; NUCLEOPLASMA (By similarity).
CC -1 SIMILARITY: BELONGS TO THE CPSP160 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF322193; AAG40326.1;
DR InterPro: IPR004871; CPSP_A.
DR Pfam: PF03178; CPSP_A; 1.
DR mRNA processing; Nuclear protein; RNA-binding.
SQ SEQUENCE 1441 AA; 160817 MW; 5D927224152AC3B9 CRC64;

Query Match
Best Local Similarity 62.5%; Score 29; DB 1; Length 1441;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLXDDLE 8
DB 1426 ILLDLE 1433

RESULT 21
CPSA_HUMAN STANDARD; PRT; 1442 AA.
AC Q10570;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-OCT-2001 (Rel. 40, Last annotation update)
DE Cleavage and polyadenylation specificity factor, 160 kDa subunit (CPSP
DE 160 kDa subunit).
GN CPSF1 OR CPSP160.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
ON NCBI_TaxID=9606;
RX MEDLINE=96067159; PubMed=7590244;
RN [1]
RP SEQUENCE FROM N.A.
RA Murthy K.G., Manley J.L.;
RT "The 160-kD subunit of human cleavage-polyadenylation specificity
RT factor coordinates pre-mRNA 3'-end formation."
RL Genes Dev. 9:2672-2683(1995).
CC -1 FUNCTION: CPSP PLAYS A KEY ROLE IN PRE-MRNA 3'-END FORMATION,
CC RECOGNIZING THE AAUAAA SIGNAL, SEQUENCE AND INTERACTING WITH
CC POLY(A) POLYMERASE AND OTHER FACTORS TO BRING ABOUT CLEAVAGE AND
CC POLY(A) ADDITION. THIS SUBUNIT IS INVOLVED IN THE RNA RECOGNITION
CC STEP OF THE POLYADENYLATION REACTION.
CC -1 SUBUNIT: CPSP IS A HETEROTETRAMER COMPOSED OF FOUR DISTINCT
CC SUBUNITS 160, 100, 70 AND 30 kDa.
CC -1 SUBCELLULAR LOCATION: NUCLEAR; nucleoplasm.
CC -1 PTM: THE N-TERMINUS IS BLOCKED.
CC -1 SIMILARITY: BELONGS TO THE CPSP160 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

```



or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).

CC EMBL: U37012; AAC50293.1; -  
 DR Genew; HGNC:2324; CPST1.  
 DR MIM: 606027; -  
 DR InterPro; IPR004871; CPST\_A.  
 DR Pfam; PF03178; CPST\_A; 1.  
 KW mRNA processing; Nuclear protein; RNA-binding.  
 FT DOMAIN 893 908 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).  
 SQ SEQUENCE 1442 AA; 160822 MW; 7FEFDEB28D7FCF8 CRC64;

Query Match  
 Best Local Similarity 76.3%; Score 29; DB 1; Length 1442;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 VLXDDLE 8  
 Db 1427 ILDDLE 1434

## RESULT 22

CPST\_BOVIN STANDARD; PRT: 1444 AA.  
 AC Q10569;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Cleavage and polyadenylation specificity factor, 160 kDa subunit (CPST  
 DE 160 kDa subunit).  
 GN CPST1 OR CPST160.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.  
 RC TISSUE=Thymus;  
 RA MEDLINE=95380277; PubMed=7651824;  
 RX Jenny A., Keller W.;  
 RT "Cloning of cDNAs encoding the 160 kDa subunit of the bovine cleavage  
 RT and polyadenylation specificity factor";  
 RL Nucleic Acids Res. 23:2629-2635(1995).  
 RN [2]  
 RP CHARACTERIZATION.  
 RX MEDLINE=92097544; PubMed=1756731;  
 RA Keller W., Bienroth S., Lang K.M., Christofori G.;  
 RT "Cleavage and polyadenylation factor CPF specifically interacts with  
 RT the pre-mRNA 3' processing signal AAUAAA";  
 RL EMBO J. 10:4241-4249(1991).

-1- FUNCTION: CPST PLAYS A KEY ROLE IN PRE-MRNA 3'-END FORMATION,  
 RECOGNIZING THE AAUAAA SIGNAL SEQUENCE AND INTERACTING WITH  
 POLY(A) POLYMERASE AND OTHER FACTORS TO BRING ABOUT CLEAVAGE AND  
 STEP OF THE POLYADENYLATION REACTION.  
 CC SUBUNIT: CPST IS A HETEROTETRAMER COMPOSED OF FOUR DISTINCT  
 CC SUBUNITS 160, 100, 70 AND 30 kDa.  
 CC -1- SUBCELLULAR LOCATION: Nuclear; nucleoplasm.  
 CC -1- PTM: THE N-TERMINUS IS BLOCKED.  
 CC -1- SIMILARITY: BELONGS TO THE CPST160 FAMILY.

CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation  
 CC at the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: X83097; CAA58152.1; -  
 DR InterPro; IPR004871; CPST\_A.  
 DR Pfam: PF03178; CPST\_A; 1.  
 KW mRNA processing; Nuclear protein; RNA-binding.

FT DOMAIN 894 909 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).  
 SQ SEQUENCE 1444 AA; 161214 MW; 226B3A4F9812E0FA CRC64;

Query Match  
 Best Local Similarity 76.3%; Score 29; DB 1; Length 1442;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 VLXDDLE 8  
 Db 1429 ILDDLE 1436

## RESULT 23

FER\_BACST STANDARD; PRT: 81 AA.  
 ID FER\_BACST  
 AC P00212;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE Ferredoxin.  
 GN FER.  
 OS Bacillus stearothermophilus.  
 OC Bacteria; Firmicutes; Bacillales; Geobacillus.  
 OX NCBI\_TaxID=1422;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=77065140; PubMed=999643;  
 RA Hase T., Ohmura N., Matsubara H., Mullinger R.N., Rao K.K.,  
 RA Hall D.O.;  
 RT "Amino acid sequence of a four-iron-four-sulphur ferredoxin isolated  
 RT from Bacillus stearothermophilus";  
 RL Biochem. J. 159:55-63(1976).  
 CC -1- FUNCTION: FERREDOXINS ARE IRON-SULFUR PROTEINS THAT TRANSFER  
 CC ELECTRONS IN A WIDE VARIETY OF METABOLIC REACTIONS.  
 CC -1- COFACTOR: BINDS 1 4FE-4S CLUSTER.  
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL TYPE FERREDOXIN FAMILY.  
 DR HSSP; P10245; 1102.  
 DR InterPro; IPR001080; 3fe4s-ferredoxin.  
 DR InterPro; IPR001450; 4fe4s-ferredoxin.  
 DR PRINTS; PR00352; 3FE4SFERDOXIN.  
 DR PROSITE; PS00198; 4FE4S\_FERREDOXIN; FALSE\_NEG.  
 KW Electron transport; Iron-sulfur; Repeat; 4fe-4s.  
 FT METAL 11 11 IRON-SULFUR (4FE-4S).  
 FT METAL 14 14 IRON-SULFUR (4FE-4S).  
 FT METAL 17 17 IRON-SULFUR (4FE-4S).  
 FT METAL 61 61 IRON-SULFUR (4FE-4S).  
 SQ SEQUENCE 81 AA; 8770 MW; C5FDE6EB193A525C CRC64;

Query Match  
 Best Local Similarity 73.7%; Score 26; DB 1; Length 81;  
 Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 1 VLXDDLE 9  
 Db 49 ILDDMDMA 57

## RESULT 24

FER\_BACST STANDARD; PRT: 81 AA.  
 ID FER\_BACST  
 AC P10245;  
 DT 01-MAR-1989 (Rel. 10, Created)  
 DT 01-MAR-1989 (Rel. 10, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Ferredoxin.  
 OS Bacillus stearothermophilus.  
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
 OX NCBI\_TaxID=1427;  
 RN [1]  
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).  
 RX MEDLINE=88172459; PubMed=3351918;  
 Fukuyma K., Nagahara Y., Tsukihara T., Katsube Y., Hase T.,



RA Matsubara H.;  
 RT "tertiary structure of Bacillus thermoproteolyticus [4Fe-4S]  
 ferredoxin. Evolutionary implications for bacterial ferredoxins.";  
 RL J. Mol. Biol. 199;183-193(1988).  
 RN [2]  
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).  
 RX MEDLINE-90096160; PubMed-2600971;  
 RA Fukuyama K., Matsubara H., Tsukihara T., Katsube Y.;  
 RT "Structure of [4Fe-4S] ferredoxin from Bacillus thermoproteolyticus  
 refined at 2.3-A resolution. Structural comparisons of bacterial  
 ferredoxins.";  
 RL J. Mol. Biol. 210:383-398(1989).  
 RN [3]  
 RP X-RAY CRYSTALLOGRAPHY (0.92 ANGSTROMS).  
 RX MEDLINE-21686305; PubMed-11827483;  
 RA Fukuyama K., Okada T., Kakuta Y., Takahashi Y.;  
 RT "Atomic resolution structures of oxidized [4Fe-4S] ferredoxin from  
 Bacillus thermoproteolyticus in two crystal forms: systematic  
 distortion of [4Fe-4S] cluster in the protein.";  
 RL J. Mol. Biol. 315:1135-1166(2002).  
 CC FUNCTION: FERREDOXINS ARE IRON-SULFUR PROTEINS THAT TRANSFER  
 ELECTRONS IN A WIDE VARIETY OF METABOLIC REACTIONS.  
 CC -1- COFACTOR: BINDS 1 4FE-4S CLUSTER.  
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL TYPE FERREDOXIN FAMILY.  
 DR PDB: 1I02; 13-FEB-02.  
 DR PDB: 1IR0; 13-FEB-02.  
 DR InterPro: IPR001080; 3FE4S-ferredoxin.  
 DR InterPro: IPR001450; 4FE4S-ferredoxin.  
 DR PRINTS: PR00352; 3FE4SFERDOXIN.  
 DR PROSITE: PS00198; 4FE4S-FERREDOXIN; FALSE-NEG.  
 KW Electron transport; Iron-sulfur; Repeat: 4FE-4S; 3D-structure.  
 FT METAL 11 11 IRON-SULFUR (4FE-4S).  
 FT METAL 14 14 IRON-SULFUR (4FE-4S).  
 FT METAL 17 17 IRON-SULFUR (4FE-4S).  
 FT METAL 61 61 IRON-SULFUR (4FE-4S).  
 FT STRAND 4 6  
 FT TURN 9 10  
 FT HELIX 17 20  
 FT TURN 22 24  
 FT TURN 25 28  
 FT STRAND 29 31  
 FT TURN 32 35  
 FT STRAND 36 39  
 FT TURN 48 59  
 FT HELIX 60 60  
 FT TURN 62 63  
 FT STRAND 66 68  
 FT TURN 77 78  
 SQ SEQUENCE 81 AA; 8770 MW; C44C3D9B193A525C CRC64;

Query Match 73.7%; Score 28; DB 1; Length 81;  
 Best Local Similarity 44.4%; Pred. NO. 32;  
 Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

OY 1 VLXDDLEA 9  
 : ||::||  
 DB 49 ILIDMDMA 57

RESULT 25  
 RL7\_RALSO STANDARD; PRT; 124 AA.  
 AC 08XU27;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DE 50S ribosomal protein L7/L12.  
 GN RPL7 OR RSC3035 OR RS04722.  
 OS Ralstonia solanacearum (Pseudomonas solanacearum).  
 OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;  
 OC Ralstonia.  
 OC NCBI\_TaxID=305;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RC STRAIN-GM11000;  
 RX MEDLINE-21681879; PubMed-11823852;  
 RA Salanoubat M., Genin S., Artiguenave F., Guzy J., Mangelot S.,  
 RA Arlet M., Billault A., Brothier P., Camus J.C., Cattolico L.,  
 RA Chandler M., Choise N., Claudel-Renard C., Cunne S., Demange N.,  
 RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schlex T.,  
 RA Sigulier P., Thebault P., Whalen M., Wincker P., Levy M.,  
 RA Weissensbach J., Boucher C.A.;  
 RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";  
 RL Nature 415:497-502(2002).  
 CC FUNCTION: Seems to be the binding site for several of the factors  
 involved in protein synthesis and appears to be essential for  
 accurate translation (by similarity).  
 CC -1- SIMILARITY: BELONGS TO THE L12P FAMILY OF RIBOSOMAL PROTEINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AL646073; CAD16744.1;  
 DR InterPro: IPR00206; Ribosomal\_L12.  
 DR Pfam: PF00542; Ribosomal\_L12; 1.  
 DR PRODOM: PD001326; Ribosomal\_L12; 1.  
 DR TRIFAMS: TIGR00855; L12; 1.  
 KW Ribosomal protein; Complete proteome.  
 SQ SEQUENCE 124 AA; 12520 MW; 271C3CF7180269E CRC64;

Query Match 73.7%; Score 28; DB 1; Length 124;  
 Best Local Similarity 55.6%; Pred. NO. 51;  
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 VLXDDLEA 9  
 : ||::||  
 DB 3 ITKDDLEA 11

RESULT 26  
 NIKR\_ECOLI STANDARD; PRT; 133 AA.  
 ID NIKR\_ECOLI  
 AC P28910; Q47559;  
 DT 01-DEC-1992 (Rel. 24, Created)  
 DT 01-DEC-1992 (Rel. 24, Last sequence update)  
 DE 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Nickel responsive regulator.  
 GN NIKR OR B3481 OR Z4873 OR ECS4348.  
 OS Escherichia coli, and  
 OS Escherichia coli O157:H7.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Escherichia.  
 OC NCBI\_TaxID=562; 83334;  
 OX NCB1  
 RN [1]  
 RP SEQUENCE FROM N.A., SEQUENCE OF 1-7, AND CHARACTERIZATION.  
 RP MEDLINE-99102235; PubMed-9882686;  
 RA De Pina K., Desjardin V., Mandrand-Berthelot M.-A., Giordano G.,  
 RA Wu L.-F.;  
 RT "Isolation and characterization of the nIKR gene encoding a nickel-  
 responsive regulator in Escherichia coli.";  
 RL J. Bacteriol. 181:670-674(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RP STRAIN-K12;  
 RX MEDLINE-93559920; PubMed-8387990;  
 RA Zhao S., Sandt C.H., Feulner G., Vlazny D.A., Gray J.A., Hill C.W.;  
 RT "Rhs elements of Escherichia coli K-12: complex composites of shared  
 RT and unique components that have different evolutionary histories.";  
 RL J. Bacteriol. 175:2799-2808(1993).  
 RN [3]  
 RP SEQUENCE FROM N.A.

RC STRAIN-K12 / MG1655;  
 RX MEDLINE-94316500; PubMed-8041620;  
 RA Sofia H.J., Burland V., Daniels D.L., Plunkett G. III, Blattner F.R.;  
 RT "Analysis of the *Escherichia coli* genome. V. DNA sequence of the  
 RL region from 76.0 to 81.5 minutes.";  
 RN Nucleic Acids Res. 22:2576-2586(1994).  
 [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-0157:H7 / ATCC 700927;  
 RX MEDLINE-21074935; PubMed-11205551;  
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,  
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,  
 RA Postai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,  
 RA Grodeck E.J., Davis N.W., Lim A., Dimantia E.T., Potamousis K.,  
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,  
 RA Welch R.A., Blattner F.R.;  
 RT "Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7.";  
 RL Nature 409:529-533(2001).  
 [5]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-0157:H7 / RIMD 0509952;  
 RX MEDLINE-21156231; PubMed-11258796;  
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,  
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,  
 RA Iida T., Takami H., Honda T., Sasakawa C., Ogawara N., Yasunaga T.,  
 RA Kubota S., Shiba T., Hattori M., Shinagawa H.;  
 RT "Complete genome sequence of enterohaemorrhagic *Escherichia coli*  
 RL O157:H7 and genomic comparison with a laboratory strain K-12.";  
 RN DNA Res. 8:11-22(2001).  
 [6]  
 RP SEQUENCE OF 101-133 FROM N.A.  
 RC STRAIN-K12 / W3110;  
 RX MEDLINE-94162733; PubMed-7764507;  
 RA Talukder A.A., Yanai S., Yamada M.;  
 RT "Analysis of products of the *Escherichia coli* genomic genes and  
 RL regulation of their expressions: an applicable procedure for genomic  
 analysis of other microorganisms.";  
 RN Biosci. Biotechnol. Biochem. 58:117-120(1994).  
 [7]  
 RP CHARACTERIZATION, AND MUTAGENESIS OF ARG-3.  
 RC STRAIN-K12 / MC1061;  
 RX MEDLINE-20061003; PubMed-10595554;  
 RA Chivers P.T., Sauer R.T.;  
 RT "NikR is a ribbon-helix-helix DNA-binding protein.";  
 RL Protein Sci. 8:2494-2500(1999).  
 [8]  
 RP CHARACTERIZATION.  
 RC STRAIN-K12 / MC1061;  
 RX MEDLINE-20347879; PubMed-10787413;  
 RA Chivers P.T., Sauer R.T.;  
 RT "Regulation of high affinity nickel uptake in bacteria. Ni2+-dependent  
 RL interaction of NikR with wild-type and mutant operator sites.";  
 RN J. Biol. Chem. 273:19735-19741(2000).  
 [9]  
 RP FUNCTION: Transcriptional repressor of the nikABCDE operon. Is  
 CC active in the presence of excessive concentrations of  
 CC intracellular nickel.  
 CC -1- COFACTOR: Binds 1 nickel ion per subunit.  
 CC -1- SUBUNIT: Homodimer.  
 CC -1- SIMILARITY: BELONGS TO THE COPG/NIKR FAMILY OF TRANSCRIPTIONAL  
 CC REGULATORS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: Y08992; CAA70150.1; -  
 CC EMBL: 102370; AAC61882.1; -  
 CC EMBL: U00039; AAB18456.1; -  
 CC EMBL: AE000423; AAC76506.1; -

DR EMBL: AE0005572; AAC58608.1; -  
 DR EMBL: AP002565; BAB37771.1; -  
 DR EMBL: D21140; BAA04676.1; ALT\_INIT.  
 DR Ecogen; EG11519; nikR.  
 DR InterPro: IPR002145; COPG\_HTH\_4.  
 DR Pfam: PF01402; HTH\_4; 1.  
 KW Transcription regulation; DNA-binding; Repressor; Nickel;  
 KM Metal-binding; Complete proteome.  
 FT METAL 76  
 FT METAL 87  
 FT METAL 89  
 FT METAL 95  
 FT MUTAGEN 3  
 FT CONFLICT 131 132  
 SQ SEQUENCE 133 AA: 15094 MW: 9495F838D3FFEF1 CRC64;  
 KE -> EGRSLDGLGVN (IN REF. 6).  
 Query Match 73.7%; Score 28; DB 1; Length 133;  
 Best Local Similarity 85.7%; Pred. No. 55;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 2 LXDILLE 8  
 Db 8 LDDILLE 14  
 RESULT 27  
 NIKR\_SALTY  
 ID NIKR\_SALTY STANDARD: PRT; 133 AA.  
 AC Q8XK6;  
 DT 15-JUN-2002 (Rel. 41, Created)  
 DT 15-JUN-2002 (Rel. 41, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Nickel responsive regulator.  
 OS NIKR OR STM3584 OR STY4227.  
 GN Salmonella typhimurium, and  
 OS Salmonella typhimurium.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Salmonella.  
 CC NCBI\_TaxID=602, 601;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES-S. typhimurium; STRAIN-LT2 / SCS0412 / ATCC 700720;  
 RX MEDLINE-21534948; PubMed-11677609;  
 RA McEliand M., Sanderson K.E., Spieth J., Clifton S.W., Latelle P.,  
 RA Courtney L., Portolillo S., Ali J., Dante M., Du F., Hou S., Layman D.,  
 RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,  
 RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,  
 RA Waterston R., Wilson R.K.;  
 RT "Complete genome sequence of *Salmonella enterica* serovar Typhimurium  
 RL LT2.";  
 RN Nature 413:852-856(2001).  
 [2]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES-S. typhimurium; STRAIN-CT18;  
 RX MEDLINE-21534947; PubMed-11677608;  
 RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,  
 RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,  
 RA Baker S., Basham D., Brooks K., Chillingworth T., Connerton P.,  
 RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,  
 RA Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jørgensen K.,  
 RA Krogh A., Larsen T.S., Leather S., Moule S., O'Goora P., Parry C.,  
 RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,  
 RA Whitehead S., Barrall B.G.;  
 RT "Complete genome sequence of a multiple drug resistant *Salmonella*  
 RL enterica serovar Typhimurium CT18.";  
 RN Nature 413:848-852(2001).  
 CC -1- FUNCTION: Transcriptional regulator (Potential).  
 CC -1- COFACTOR: Binds 1 nickel ion per subunit (By similarity).  
 CC -1- SUBUNIT: Homodimer (By similarity).  
 CC -1- SIMILARITY: BELONGS TO THE COPG/NIKR FAMILY OF TRANSCRIPTIONAL  
 CC REGULATORS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AL627281: CAD08046.1: -  
 DR Stycene; SG77777.nhr.  
 DR InterPro: IPR002145: COGc\_HTH\_4.  
 DR Pfam: PF01402: HTH\_4: 1.  
 DR Transcription regulation; DNA-binding; Nickel; Metal-binding;  
 KW Complete proteome.  
 FT METAL 76 NICKEL (POTENTIAL).  
 FT METAL 87 NICKEL (POTENTIAL).  
 FT METAL 89 NICKEL (POTENTIAL).  
 FT METAL 95 NICKEL (POTENTIAL).  
 SQ SEQUENCE 133 AA; 15065 MW; FE03D857053E2424 CRC64;  
 Query Match 73.7%; Score 28; DB 1; Length 133;  
 Best Local Similarity 85.7%; Pred. No. 55;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 2 LXDLE 8  
 DB 8 LDDLE 14  
 RESULT 28  
 PKSI\_BACSU STANDARD; PRT; 249 AA.  
 ID PKSI\_BACSU  
 AC P40802;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Putative polyketide biosynthesis enoyl-CoA hydratase homolog pksI.  
 GN PKSI.  
 OS Bacillus subtilis.  
 CC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
 CC NCBI\_TaxID=1423;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168 / P81424;  
 RX MEDLINE=98044033; PubMed=9384377;  
 RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,  
 Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,  
 Bortles R., Boursier L., Brans A., Braun M., Brijnelli S.C., Bron S.,  
 Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,  
 Choi S.K., Codani U.J., Connerton I.F., Cummings N.J., Daniel R.A.,  
 Denzot F., Devant K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,  
 Entlan K.D., Errington J., Fabret C., Ferrari E., Foulger D.,  
 Fritz C., Fujita M., Fujita Y., Fuma S., Gallizi A., Galleron N.,  
 Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grand G.,  
 Guisepi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,  
 Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,  
 Joris B., Karamata D., Kashiwara Y., Klauer-Blanchard M., Klein C.,  
 Kurihara K., Kurihara Y., Kurihara Y., Kurihara Y., Kurihara Y.,  
 Lee S.M., Levine A., Liu H., Masuda S., Maue C., Medigue C.,  
 Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,  
 Noone D., O'Reilly M., Ogawa K., Ogilvie A., Oudega B., Park S.H.,  
 Parro V., Pohl T.M., Portetlelle D., Portolillo S., Prescott A.M.,  
 Prescan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,  
 Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadate Y.,

RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,  
 Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,  
 Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,  
 Takeuchi M., Tanakoshi A., Tanaka T., Terpstra P., Tognoni A.,  
 Toosto V., Uchiyama S., Vandenbol M., Vanier F., Vasseroiti A.,  
 Viati A., Wambuit R., Wedler H., Welteregger T.,  
 Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,  
 Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.,  
 RA "the complete genome sequence of the Gram-positive bacterium Bacillus  
 RT subtilis".  
 RL Nature 390:249-256(1997).  
 CC -1- FUNCTION: POTENTIALLY INVOLVED IN SOME INTERMEDIATE STEPS FOR  
 CC THE SYNTHESIS OF A POLYKETIDE MOLECULE WHICH MAY BE INVOLVED IN  
 CC SECONDARY METABOLISM.  
 CC -1- SIMILARITY: BELONGS TO THE ENOYL-COA HYDRATASE/ISOMERASE FAMILY.  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: U11039: AAA85142.1: -  
 DR EMBL: 299112: CAB13588.1: -  
 DR EMBL: 299113: CAB13599.1: -  
 DR Subtilisin; BG10928: pksI.  
 DR InterPro: IPR001753: ENCOA\_hydratase.  
 DR Pfam: PF00378: Ech, 1.  
 DR PROSITE: PS00166: ENOYL\_COA\_HYDRATASE; 1.  
 DR Antibiotic biosynthesis; Lyase; Complete proteome.  
 KW SEQUENCE 249 AA; 27954 MW; FD7CDA0DCFD915 CRC64;  
 Query Match 73.7%; Score 28; DB 1; Length 249;  
 Best Local Similarity 62.5%; Pred. No. 11e+02;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 QY 2 LXDLE 9  
 DB 33 LDDLE 40  
 RESULT 29  
 VMT1\_TABAN STANDARD; PRT; 252 AA.  
 ID VMT1\_TABAN  
 AC P03487;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE Matrix protein M1.  
 OS Influenza A virus (strain A/Bangkok/1/79).  
 CC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;  
 CC Influenza A viruses; Influenzavirus A.  
 CC NCBI\_TaxID=11325;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=84005895; PubMed=6688599;  
 RX Ortin J., Martinez C., del Rio L., Davila M., Lopez-Galindez C.,  
 Villanueva N., Domingo E.;  
 RA "Evolution of the nucleotide sequence of influenza virus RNA segment  
 RT 7 during drift of the H3N2 subtype";  
 RL Gene 23:233-239(1987).  
 CC -1- FUNCTION: THIS PROTEIN FORMS A CONTINUOUS SHELL ON THE INNER SIDE  
 CC OF THE LIPID BILAYER. MAINTAINS THE STRUCTURAL INTEGRITY OF THE  
 CC VIRUS PARTICLE THROUGH HYDROPHOBIC INTERACTIONS. ALSO MEDIATES  
 CC THE ENCAPSULATION OF RNA NUCLEOPROTEIN CORES INTO THE MEMBRANE  
 CC ENVELOPE.  
 CC -1- ALTERNATIVE PRODUCTS: MATRIX PROTEINS M1 AND M2 ARE PRODUCED BY  
 CC ALTERNATIVE SPLICING OF RNA SEGMENT 7. ONLY THE FIRST 9 RESIDUES  
 CC ARE SHARED BY THE TWO PROTEINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: K01140; AAA43092.1; -  
 DR PIR: A04080; MFLV1K.  
 DR HSSP: P03485; 1AA7.  
 DR InterPro: IPR001561; Flu\_M1.  
 DR Pfam: PF00598; Flu\_M1.1.  
 DR ProDom: PD001061; Flu\_M1.1.  
 DR Matrix protein: RNA-binding; Alternative splicing.  
 SO SEQUENCE 252 AA: 27810 MW; F892297A838466C2 CRC64;

Query Match 73.7%; Score 28; DB 1; Length 252;  
 Best Local Similarity 85.7%; Pred. No. 1.le+02;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 LKDDLE 8  
 DB 229 LKDDLE 235

RESULT 30  
 ID VMT1\_IACKB STANDARD; PRT; 252 AA.  
 AC P36347;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE Matrix protein M1.  
 OS Influenza A virus (strain A/Chicken/Brescia/1902).  
 CC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;  
 CC Influenza A viruses; Influenzavirus A.  
 CC NCBI\_TaxID=36418;  
 CC [1]  
 CC SEQUENCE FROM N.A.  
 CC MEDLINE=92109567; PubMed=1530908;  
 CC Klimov A., Proesch S., Schaefer J., Bucher D.;  
 CC "Subtype H7 influenza viruses: comparative antigenic and molecular  
 CC analysis of the HA-, M-, and NS-genes";  
 CC Arch. Virol. 122:143-161(1992).  
 CC -!- FUNCTION: THIS PROTEIN FORMS A CONTINUOUS SHELL ON THE INNER SIDE  
 CC OF THE LIPID BILAYER. MAINTAINS THE STRUCTURAL INTEGRITY OF THE  
 CC VIRUS PARTICLE THROUGH HYDROPHOBIC INTERACTIONS. ALSO MEDIATES  
 CC THE ENCAPSIDATION OF RNA NUCLEOPROTEIN CORES INTO THE MEMBRANE  
 CC ENVELOPE.  
 CC -!- ALTERNATIVE PRODUCTS: MATRIX PROTEINS M1 AND M2 ARE PRODUCED BY  
 CC ALTERNATIVE SPLICING OF RNA SEGMENT 7. ONLY THE FIRST 9 RESIDUES  
 CC ARE SHARED BY THE TWO PROTEINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: L37795; AAA56804.1; -  
 DR PIR: B45539; B45539.  
 DR HSSP: P03485; 1AA7.  
 DR InterPro: IPR001561; Flu\_M1.  
 DR Pfam: PF00598; Flu\_M1.1.  
 DR ProDom: PD001061; Flu\_M1.1.  
 DR Matrix protein: RNA-binding; Alternative splicing.  
 SO SEQUENCE 252 AA: 27964 MW; F4091A18C48C9800 CRC64;

Query Match 73.7%; Score 28; DB 1; Length 252;  
 Best Local Similarity 85.7%; Pred. No. 1.le+02;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 LKDDLE 8  
 DB 229 LKDDLE 235

RESULT 31  
 ID VMT1\_IACFW STANDARD; PRT; 252 AA.  
 AC P10918;  
 DT 01-JUL-1989 (Rel. 11, Created)  
 DT 01-JUL-1989 (Rel. 11, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE Matrix protein M1.  
 OS Influenza A virus (strain A/Fort Warren/1/50), and  
 CC Influenza A virus (strain A/Singapore/1/57).  
 CC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;  
 CC Influenza A viruses; Influenzavirus A.  
 CC NCBI\_TaxID=11381, 11463;  
 CC [1]  
 CC SEQUENCE FROM N.A.  
 CC MEDLINE=89240054; PubMed=2701939;  
 CC Zebadee S.L., Lamb R.A.;  
 CC "Nucleotide sequences of Influenza A virus RNA segment 7: a  
 CC comparison of five isolates";  
 CC Nucleic Acids Res. 17:2870-2870(1989).  
 CC -!- FUNCTION: THIS PROTEIN FORMS A CONTINUOUS SHELL ON THE INNER SIDE  
 CC OF THE LIPID BILAYER. MAINTAINS THE STRUCTURAL INTEGRITY OF THE  
 CC VIRUS PARTICLE THROUGH HYDROPHOBIC INTERACTIONS. ALSO MEDIATES  
 CC THE ENCAPSIDATION OF RNA NUCLEOPROTEIN CORES INTO THE MEMBRANE  
 CC ENVELOPE.  
 CC -!- ALTERNATIVE PRODUCTS: MATRIX PROTEINS M1 AND M2 ARE PRODUCED BY  
 CC ALTERNATIVE SPLICING OF RNA SEGMENT 7. ONLY THE FIRST 9 RESIDUES  
 CC ARE SHARED BY THE TWO PROTEINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: X08091; CAA30888.1; -  
 DR EMBL: X08093; CAA30892.1; -  
 DR PIR: S04056; S04056.  
 DR PIR: S04060; S04060.  
 DR PIR: S14616; S14616.  
 DR HSSP: P03485; 1AA7.  
 DR InterPro: IPR001561; Flu\_M1.  
 DR Pfam: PF00598; Flu\_M1.1.  
 DR ProDom: PD001061; Flu\_M1.1.  
 DR Matrix protein: RNA-binding; Alternative splicing.  
 SO SEQUENCE 252 AA: 27903 MW; 7317397B86DFC877 CRC64;

Query Match 73.7%; Score 28; DB 1; Length 252;  
 Best Local Similarity 85.7%; Pred. No. 1.le+02;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 LKDDLE 8  
 DB 229 LKDDLE 235

RESULT 32  
 ID VMT1\_IAPFR STANDARD; PRT; 252 AA.  
 AC P03488;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE Matrix protein M1.  
 OS Influenza A virus (strain A/Fowl plague virus/Rostock/34).

```

OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC Influenza A viruses; Influenzavirus A.
OC NCBI_TaxID=11383;
RX [1]
RN SEQUENCE FROM N.A.
RP MEDLINE=83058699; PubMed=6292344;
RA McCauley J.W., Mahy B.W.J., Ingalls S.C.;
RT "Nucleotide sequence of fowl plague virus RNA segment 7.";
RL J. Gen. Virol. 58:211-215(1982).
CC -1- FUNCTION: THIS PROTEIN FORMS A CONTINUOUS SHELL ON THE INNER SIDE
CC OF THE LIPID BILAYER. MAINTAINS THE STRUCTURAL INTEGRITY OF THE
CC VIRUS PARTICLE THROUGH HYDROPHOBIC INTERACTIONS. ALSO MEDIATES
CC THE ENCAPSIDATION OF RNA NUCLEOPROTEIN CORES INTO THE MEMBRANE
CC ENVELOPE.
CC -1- ALTERNATIVE PRODUCTS: MATRIX PROTEINS M1 AND M2 ARE PRODUCED BY
CC ALTERNATIVE SPLICING OF RNA SEGMENT 7. ONLY THE FIRST 9 RESIDUES
CC ARE SHARED BY THE TWO PROTEINS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; X05905; CAA2934.1; -.
DR PIR; A04081; MFTVLF.
DR HSSP; P03485; 1AA7.
DR InterPro: IPR001561; FLU_M1.
DR Pfam; PF00598; FLU_M1; 1.
DR ProDom: PD001061; FLU_M1; 1.
DR Matrix protein; RNA-binding; Alternative splicing.
SQ SEQUENCE 252 AA; 27925 MW; 6E9B924276480FC4 CRC64;

OY 2 LKDDLE 8
Db 229 LKDDLE 235

Query Match 73.7%; Score 28; DB 1; Length 252;
Best Local Similarity 85.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

RESULT 33
VT1_LAFPM STANDARD; PRT; 252 AA.
ID VT1_LAFPM
AC P05775;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Matrix protein M1.
OS Influenza A virus (strain A/Fowl plague virus/Meybridge).
OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC Influenza A viruses; Influenzavirus A.
OC NCBI_TaxID=11384;
RX [1]
RN SEQUENCE FROM N.A.
RP MEDLINE=88223193; PubMed=3414185;
RA Markushin S., Ghlasi H., Sokolov N., Shilov A., Sintsin B., Brown D.,
RA Klimov A., Nayak D.;
RT "Nucleotide sequence of RNA segment 7 and the predicted amino
RT sequence of M1 and M2 proteins of FpV/Meybridge (H7N7) and WSN (H1N1)
RT influenza viruses.";
RL Virus Res. 10:263-272(1988).
RN [2]
RN SEQUENCE FROM N.A.
RP MEDLINE=88221675; PubMed=3450279;
RA Karginov V.A., Blinov V.M., Salironov P.F., Mamayev L.V., Golovin S.Y.,
RA Netesov S.V., Samokhvalov E.I., Sharova N.K., Yufarov V.P.,
RA Yufarov L.V., Bukrinskaya A.G.;
RT "Comparative analysis of primary structure of M-genes in remantidine-
RT resistant and remantidine-sensitive strains of influenza virus

```

```

RT A/FpV/Meybridge (H7N7) strains.";
RL Bioorg. Khim. 13:1638-1643(1987).
CC -1- FUNCTION: THIS PROTEIN FORMS A CONTINUOUS SHELL ON THE INNER SIDE
CC OF THE LIPID BILAYER. MAINTAINS THE STRUCTURAL INTEGRITY OF THE
CC VIRUS PARTICLE THROUGH HYDROPHOBIC INTERACTIONS. ALSO MEDIATES
CC THE ENCAPSIDATION OF RNA NUCLEOPROTEIN CORES INTO THE MEMBRANE
CC ENVELOPE.
CC -1- ALTERNATIVE PRODUCTS: MATRIX PROTEINS M1 AND M2 ARE PRODUCED BY
CC ALTERNATIVE SPLICING OF RNA SEGMENT 7. ONLY THE FIRST 9 RESIDUES
CC ARE SHARED BY THE TWO PROTEINS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; M2917; AAA43251.1; -.
DR PIR; M38299; AAA43314.1; ALT_TERM.
DR PIR; PNO083; PNO083.
DR PIR; PNO086; PNO086.
DR PIR; S07945; S07945.
DR HSSP; P03485; 1AA7.
DR InterPro: IPR001561; FLU_M1.
DR Pfam; PF00598; FLU_M1; 1.
DR ProDom: PD001061; FLU_M1; 1.
DR Matrix protein; RNA-binding; Alternative splicing.
KM VARIANT 109 109
FT CONFLICT 28 28
FT CONFLICT 41 41
FT CONFLICT 53 53
FT CONFLICT 95 95
FT CONFLICT 101 101
FT CONFLICT 110 110
FT CONFLICT 144 144
FT CONFLICT 158 158
FT CONFLICT 181 181
FT CONFLICT 212 213
SQ SEQUENCE 252 AA; 27904 MW; 8B823F78A2E91573 CRC64;

OY 2 LKDDLE 8
Db 229 LKDDLE 235

Query Match 73.7%; Score 28; DB 1; Length 252;
Best Local Similarity 85.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

RESULT 34
VT1_IAMAN STANDARD; PRT; 252 AA.
ID VT1_IAMAN
AC P08381;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Matrix protein M1.
OS Influenza A virus (strain A/Mallard/New York/6750/78).
OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC Influenza A viruses; Influenzavirus A.
OC NCBI_TaxID=11435;
RX [1]
RN SEQUENCE FROM N.A.
RP MEDLINE=86115422; PubMed=3080604;
RA Buckler-White A.J., Naeye C.W., Murphy B.R.;
RT "Characterization of a gene coding for M proteins which is involved
RT in host range restriction of an avian influenza A virus in monkeys.";
RL J. Virol. 57:697-700(1986).
CC -1- FUNCTION: THIS PROTEIN FORMS A CONTINUOUS SHELL ON THE INNER SIDE
CC OF THE LIPID BILAYER. MAINTAINS THE STRUCTURAL INTEGRITY OF THE
CC VIRUS PARTICLE THROUGH HYDROPHOBIC INTERACTIONS. ALSO MEDIATES

```

THE ENCAPSULATION OF RNA NUCLEOPROTEIN CORES INTO THE MEMBRANE ENVELOPE.

CC -1- ALTERNATIVE PRODUCTS: MATRIX PROTEINS M1 AND M2 ARE PRODUCED BY ALTERNATIVE SPLICING OF RNA SEGMENT 7. ONLY THE FIRST 9 RESIDUES ARE SHARED BY THE TWO PROTEINS.

-----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: M12699; AAA43313.1; -

DR PIR: A29511; MFTVM.

DR HSSP: P03485; 1AA7.

DR InterPro: IPR001561; FLU\_M1.

DR Pfam: PF00598; FLU\_M1.1.

DR ProDom: PD001061; FLU\_M1.1.

KW Matrix protein; RNA-binding; Alternative splicing.

SQ SEQUENCE 252 AA; 27926 MW; A9E238FAC2044A45 CRC64;

Query Match 73.7%; Score 28; DB 1; Length 252;  
Best Local Similarity 85.7%; Pred. No. 1.1e+02;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 LKDDLE 8  
DB 229 LKDDLE 235

RESULT 35

VM11\_TAU0

ID VM11\_TAU0 STANDARD; PRT; 252 AA.

AC P03486; P10919;

DT 21-JUL-1986 (Rel. 01, Created)

DT 01-FEB-1996 (Rel. 33, Last sequence update)

DT 15-JUL-1998 (Rel. 36, Last annotation update)

DE Matrix protein M1.

OS Influenza A virus (strain A/Udorn/307/72), and

OC Influenza A virus (strain A/Port Chalmers/1/73).

OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;

OC Influenza A viruses; Influenzavirus A.

OX NCBI\_TaxID=11480; 11454;

RM [1]

RP SEQUENCE OF 1-13 AND 210-252 FROM N.A.

RC STRAIN-A/Udorn/307/72;

RX MEDLINE=82037801; PubMed=6945577;

RA Lamb R.A., Lai C.-J., Choppin P.W.;

RT "Sequences of mRNAs derived from genome RNA segment 7 of influenza virus: colinear and interrupted mRNAs code for overlapping proteins.";

RT Proc. Natl. Acad. Sci. U.S.A. 78:4170-4174(1981).

RL [2]

RP SEQUENCE FROM N.A.

RC STRAIN-A/Udorn/307/72;

RX MEDLINE=81251059; PubMed=7257189;

RA Lamb R.A., Lai C.-J.;

RT "Conservation of the influenza virus membrane protein (M1) amino acid sequence and an open reading frame of RNA segment 7 encoding a second protein (M2) in H1N1 and H3N2 strains.";

RT Virology 112:746-751(1981).

RL [3]

RP SEQUENCE FROM N.A.

RC STRAIN-A/Port Chalmers/1/73;

RX MEDLINE=69240054; PubMed=2701939;

RA Zebedee S.L., Lamb R.A.;

RT "Nucleotide sequences of influenza A virus RNA segment 7: a comparison of live isolates.";

RT Nucleic Acids Res. 17:2870-2870(1989).

CC -1- FUNCTION: THIS PROTEIN FORMS A CONTINUOUS SHELL ON THE INNER SIDE OF THE LIPID BILAYER. MAINTAINS THE STRUCTURAL INTEGRITY OF THE

VIRUS PARTICLE THROUGH HYDROPHOBIC INTERACTIONS. ALSO MEDIATES THE ENCAPSULATION OF RNA NUCLEOPROTEIN CORES INTO THE MEMBRANE ENVELOPE.

CC -1- ALTERNATIVE PRODUCTS: MATRIX PROTEINS M1 AND M2 ARE PRODUCED BY ALTERNATIVE SPLICING OF RNA SEGMENT 7. ONLY THE FIRST 9 RESIDUES ARE SHARED BY THE TWO PROTEINS.

-----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: J02167; AAA43304.1; -

DR EMBL: X08092; CA30880.1; -

DR EMBL: X08090; CA30886.1; -

DR PIR: A94326; MFTVC.

DR PIR: A94058; S04058.

DR PIR: S04054; S04054.

DR HSSP: P03485; 1AA7.

DR InterPro: IPR001561; FLU\_M1.

DR Pfam: PF00598; FLU\_M1.1.

DR ProDom: PD001061; FLU\_M1.1.

KW Matrix protein; RNA-binding; Alternative splicing.

FT VARIANT 41 41

SQ SEQUENCE 252 AA; 27804 MW; 92522D3E87DAC3C6 CRC64;

Query Match 73.7%; Score 28; DB 1; Length 252;  
Best Local Similarity 85.7%; Pred. No. 1.1e+02;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 LKDDLE 8  
DB 229 LKDDLE 235

RESULT 36

VM11\_TAM1L

ID VM11\_TAM1L STANDARD; PRT; 252 AA.

AC P05777;

DT 01-NOV-1988 (Rel. 09, Created)

DT 01-JUL-1989 (Rel. 11, Last sequence update)

DT 15-JUL-1998 (Rel. 36, Last annotation update)

DE Matrix protein M1.

OS Influenza A virus (strain A/Wilson-Smith/33).

OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;

OC Influenza A viruses; Influenzavirus A.

OX NCBI\_TaxID=11487;

RM [1]

RP SEQUENCE FROM N.A.

RC MEDLINE=89240054; PubMed=2701939;

RA Zebedee S.L., Lamb R.A.;

RT "Nucleotide sequences of influenza A virus RNA segment 7: a comparison of live isolates.";

RT Nucleic Acids Res. 17:2870-2870(1989).

RL [2]

RP SEQUENCE FROM N.A.

RC MEDLINE=88323193; PubMed=3414185;

RA Markushin S., Ghiasi H., Sokolov N., Shilov A., Sinitin B., Brown D., Klimov A., Nayak D.;

RT "Nucleotide sequence of RNA segment 7 and the predicted amino acid sequence of M1 and M2 proteins of FPV/Meybridge (H7N7) and WSN (H1N1) influenza viruses.";

RT Virus Res. 10:263-272(1988).

RL [3]

RP SEQUENCE FROM N.A.

RC MEDLINE=88179567; PubMed=3354209;

RA Baylors N.W., Zhiping Y.L., Wagner R.R.;

RT "Transient expression and sequence of the matrix (M1) gene of WSN influenza A virus in a vaccinia vector.";

RT Virology 163:618-621(1988).

```

CC -1- FUNCTION: THIS PROTEIN FORMS A CONTINUOUS SHELL ON THE INNER SIDE
CC OF THE LIPID BILAYER, MAINTAINS THE STRUCTURAL INTEGRITY OF THE
CC VIRUS PARTICLE THROUGH HYDROPHOBIC INTERACTIONS. ALSO MEDIATES
CC THE ENCAPSIDATION OF RNA NUCLEOPROTEIN CORES INTO THE MEMBRANE
CC ENVELOPE.
CC -1- ALTERNATIVE PRODUCTS: MATRIX PROTEINS M1 AND M2 ARE PRODUCED BY
CC ALTERNATIVE SPLICING OF RNA SEGMENT 7. ONLY THE FIRST 9 RESIDUES
CC ARE SHARED BY THE TWO PROTEINS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; X08088; CAA30882.1; -
CC EMBL; M23920; AAA43252.1; ALT_SEQ.
CC EMBL; M19374; AAA43352.1; -
CC EMBL; L25818; AAA91325.1; -
CC PIR; A28608; MFIYMS.
CC PIR; S04050; S04050.
CC PIR; S07429; S07429.
CC HSSP; P03485; 1AA7.
CC InterPro: IPR001561; Flu_M1.
CC Pfam: PF00598; Flu_M1; 1.
CC ProDom: PD001061; Flu_M1; 1.
CC Matrix protein; RNA-binding; Alternative splicing.
CC CONFLICT 117 117 L -> F (IN REF. 3).
CC CONFLICT 219 219 I -> V (IN REF. 3).
CC CONFLICT 231 231 D -> S (IN REF. 2).
CC SEQUENCE 252 AA; 27864 MW; 5F300F18D75BBD3 CRC64;

Query Match 73.7%; Score 28; DB 1; Length 252;
Best Local Similarity 85.7%; Pred. No. 1,le+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 LKDDLE 8
Db 229 LKDDLE 235

RESULT 37
VMTL IAZ11 STANDARD; PRT; 252 AA.
AC P05776;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Matrix protein M1.
OS Influenza A virus (strain A/Swine/Iowa/15/30).
OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC Influenza A viruses; Influenzavirus A.
OX NCBI_TaxID=11500;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85042103; PubMed=6495656;
RA Nakajima K., Nobusawa E., Nakajima S.;
RT "Genetic relatedness between A/Swine/Iowa/15/30(H1N1) and human
RT Influenza viruses.";
RL Virology 139:194-198(1984).
CC -1- FUNCTION: THIS PROTEIN FORMS A CONTINUOUS SHELL ON THE INNER SIDE
CC OF THE LIPID BILAYER, MAINTAINS THE STRUCTURAL INTEGRITY OF THE
CC VIRUS PARTICLE THROUGH HYDROPHOBIC INTERACTIONS. ALSO MEDIATES
CC THE ENCAPSIDATION OF RNA NUCLEOPROTEIN CORES INTO THE MEMBRANE
CC ENVELOPE.
CC -1- ALTERNATIVE PRODUCTS: MATRIX PROTEINS M1 AND M2 ARE PRODUCED BY
CC ALTERNATIVE SPLICING OF RNA SEGMENT 7. ONLY THE FIRST 9 RESIDUES
CC ARE SHARED BY THE TWO PROTEINS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

```

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; M33045; AAA43682.1; ALT_SEQ.
CC EMBL; P03485; 1AA7.
CC HSSP; P03485; 1AA7.
CC InterPro: IPR001561; Flu_M1.
CC Pfam: PF00598; Flu_M1; 1.
CC ProDom: PD001061; Flu_M1; 1.
CC Matrix protein; RNA-binding; Alternative splicing.
CC SEQUENCE 252 AA; 27630 MW; 417250718DF983D CRC64;

Query Match 73.7%; Score 28; DB 1; Length 252;
Best Local Similarity 85.7%; Pred. No. 1,le+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 LKDDLE 8
Db 229 LKDDLE 235

RESULT 38
ECHO MYCLE STANDARD; PRT; 257 AA.
AC O07137;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable enoyl-CoA hydratase echA8 (EC 4.2.1.17).
DE ECHA8 OR ML2402 OR MLCB1306.05C.
OS Mycobacterium leprae.
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1769;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TN;
RX MEDLINE=21128732; PubMed=11234002;
RA Cole S.T., Eiglmeyer K., Parkhill J., James K.D., Thomson N.R.,
RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,
RA Mungall K., Basham D., Brown D., Chillingworth T., Connor A.,
RA Davies R.M., Devlin K., Duthoy S., Bellwell T., Fraser A., Hamlin N.,
RA Holroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,
RA Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA Barrall B.G.;
RT "Massive gene decay in the leprosy bacillus.";
RL Nature 409:1007-1011(2001).
CC -1- FUNCTION: COULD POSSIBLY OXIDIZES FATTY ACIDS USING SPECIFIC
CC COMPONENTS (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: (3S)-3-hydroxyacyl-CoA = trans-2(or 3)-enoyl-
CC CoA + H(2)O.
CC -1- SIMILARITY: BELONGS TO THE ENOYL-COA HYDRATASE/ISOMERASE FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; Y13803; CAA74130.1; -
CC EMBL; AL583925; CAC31918.1; -
CC HSSP; P14604; 2DDB.
CC Lepidoma; ML2402; -
CC InterPro: IPR001753; ENCOA_hydrase.
CC Pfam: PF00378; ECH; 1.
CC PROSITE: PS00166; ENOYL_COA_HYDRATASE; 1.
CC Fatty acid metabolism; Lyase; Complete proteome.

```



SQ SEQUENCE 257 AA; 27516 MW; 25B801EB690BCD00 CRC64;

Query Match 73.7%; Score 28; DB 1; Length 257;  
Best Local Similarity 85.7%; Pred. No. 1.1e+02;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 VLXDDL 7  
11111  
Db 180 VLADDL 186

RESULT 39  
F034\_YEAST

STANDARD; PRT; 282 AA.

AC P32907;  
DT 01-OCT-1993 (Rel. 27, Created)  
DT 01-OCT-1993 (Rel. 27, Last sequence update)  
DT 01-OCT-1996 (Rel. 34, Last annotation update)  
DE F034 protein.  
GN F034 OR YNR002C OR N2029.  
OS Saccharomyces cerevisiae (Baker's Yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
OX NCBI\_TaxID=4932;  
[1]  
RN SEQUENCE FROM N.A.  
RP STRAIN=GRF88;  
RX MEDLINE=93016077; PubMed=1400451;  
RA Stettler S., Mariotte S., Riva M., Sentenac A., Thuriaux P.;  
RT "An essential and specific subunit of RNA polymerase III (C) is  
RL J. Biol. Chem. 267:21390-21395(1992).  
[2]  
RN SEQUENCE FROM N.A.  
RP STRAIN=GRF88;  
RX MEDLINE=95028151; PubMed=7941739;  
RA Lalo D., Stettler S., Mariotte S., Gendreau E., Thuriaux P.;  
RT "Organization of the centromeric region of chromosome XIV in  
RL Saccharomyces cerevisiae.";  
J. yeast 10:523-533(1994).  
[3]  
RN SEQUENCE FROM N.A.  
RP STRAIN=5288C / FY1679;  
RX MEDLINE=95208356; PubMed=7900425;  
RA Verhasselt P., Aert R., Voet M., Volckaert G.;  
RT "Twelve open reading frames revealed in the 23.6 kb segment flanking  
RT the centromere on the Saccharomyces cerevisiae chromosome XIV right  
arm.";  
J. yeast 10:1355-1361(1994).  
[4]  
RN SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
[5]  
CC -1- SIMILARITY: BELONGS TO THE GPR1/FUN34/YAAH FAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL: X63746; CAA5279.1; -  
CC EMBL: X73395; CAA5457.1; -  
CC EMBL: Z71617; CAA96278.1; -  
CC PIR: S31258; S31258.  
CC SGD: S0005285; FUN34.  
CC InterPro: IPR000791; GPR1\_Fun34\_YAAH.  
CC Pfam: PF01184; GPR1\_Fun34\_YAAH.1.  
CC ProDom: PD010188; GPR1\_Fun34\_YAAH.1.  
CC ProSite: PS01114; GPR1\_Fun34\_YAAH.1.  
CC Transmembrane.  
KW TRANSMEM 87 107 POTENTIAL.  
FT TRANSMEM 120 140 POTENTIAL.  
FT TRANSMEM 151 171 POTENTIAL.

FT TRANSMEM 185 205 POTENTIAL.

FT TRANSMEM 208 228 POTENTIAL.

FT TRANSMEM 239 259 POTENTIAL.

SQ SEQUENCE 282 AA; 30701 MW; F5E20F9324CEB199 CRC64;

Query Match 73.7%; Score 28; DB 1; Length 282;  
Best Local Similarity 75.0%; Pred. No. 1.2e+02;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 LXDLLFA 9  
11111  
Db 62 LRDDLFA 69

RESULT 40  
Y32K\_BNYVG

STANDARD; PRT; 282 AA.

AC P19231;  
DT 01-NOV-1990 (Rel. 16, Created)  
DT 01-NOV-1990 (Rel. 16, Last sequence update)  
DT 01-NOV-1990 (Rel. 16, Last annotation update)  
DE RNA-4 hypothetical 31.9 kDa protein.  
OS Beet necrotic yellow vein mosaic virus (isolate G1).  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Benyvirus.  
OX NCBI\_TaxID=12257;  
[1]  
RN SEQUENCE FROM N.A.  
RP Bouzouda S., Guille H., Jonard G., Richards K., Putz C.;  
RT "Nucleotide sequence analysis of RNA-3 and RNA-4 of beet necrotic  
RT yellow vein virus, isolates F2 and G1.";  
J. Gen. Virol. 66:1553-1564(1985).  
[2]  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL: M36897; AAA42800.1; -  
CC PIR: C44503; C44503.  
CC Hypothetical protein.  
KW SEQUENCE 282 AA; 31869 MW; AA7C0351C54FE0CC CRC64;  
SQ

Query Match 73.7%; Score 28; DB 1; Length 282;  
Best Local Similarity 66.7%; Pred. No. 1.2e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 VLXDDLFA 9  
11111  
Db 121 VLSDLECA 129

RESULT 41  
G3P\_METUA

STANDARD; PRT; 343 AA.

AC 058546;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 15-JUL-2002 (Rel. 41, Last annotation update)  
DE Glyceraldehyde 3-phosphate dehydrogenase (PC 1.2.1.59) (GAPDH)  
DE (NAD(P)-dependent glyceraldehyde-3-phosphate dehydrogenase).  
GN GAP OR Mj1146.  
OS Methanococcus jannaschii.  
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;  
OC Methanocaldococcaceae; Methanocaldococcus.  
OX NCBI\_TaxID=2190;  
[1]  
RN SEQUENCE FROM N.A.  
RP STRAIN=JAL-1 / DSM 2661 / ATCC 43067;  
RX MEDLINE=96337999; PubMed=6688087;  
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,



RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,  
 RA Keriavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,  
 RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,  
 RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,  
 RA Uterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,  
 RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,  
 RA Kleink H.-P., Fraser C.M., Smith H.O., Moese C.R., Venter J.C.;  
 RA "Complete genome sequence of the methanogenic archaeon, Methanococcus  
 RT jamaenschi.";  
 RT Science 273:1058-1073(1996).  
 CC -1- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate + phosphate +  
 CC NAD(P)(+) = 3-phospho-D-glyceroyl phosphate + NAD(P)H.  
 CC -1- PATHWAY: Second phase of glycolysis: first step.  
 CC -1- SUBUNIT: Homotrimer (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -1- SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE  
 CC DEHYDROGENASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: U67557; AAB99147.1; -  
 DR HSSP: P10618; ICF2.  
 DR TIGR: W01146; -  
 DR InterPro: IPR000173; GAP\_dhdrogenase.  
 DR Pfam: PF00044; gpdh.C.1.  
 DR PROSITE: PS00071; GAPDH: 1.  
 DR GlycoLysis: Oxidoreductase; NAD: NADP: Complete proteome.  
 FT BINDING 144 144 GLYCERALDEHYDE 3-PHOSPHATE (By  
 FT SIMILARITY).  
 SQ SEQUENCE 343 AA; 38102 MW; 81EB5810A9C838C5 CRC64;  
 Query Match 73.7%; Score 28; DB 1; Length 343;  
 Best Local Similarity 66.7%; Pred. No. 1.5e+02;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

RT "Whole-genome random sequencing and assembly of Haemophilus influenzae  
 RT Rd.";  
 RL Science 269:496-512(1995).  
 CC -1- FUNCTION: Probably plays an important role in intracellular  
 CC peptide degradation. Hydrolyzes peptides with N-terminal acidic  
 CC residues. The preferred substrates are peptides with N-terminal  
 CC Asp or Glu residues (By similarity).  
 CC -1- COFACTOR: Manganese (By similarity).  
 CC -1- SUBUNIT: Homohexamer (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M17.  
 CC -1- CAUTION: Ref.1 sequence differs from that shown due to a  
 CC frameshift in position 340. This may be a natural frameshift and  
 CC this protein could be a pseudogene.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: U32769; -; NOT\_ANNOTATED\_CDS.  
 DR TIGR: H10875; -  
 DR InterPro: IPR000819; Peptidase\_M17.  
 DR Pfam: PF00883; Peptidase\_M17.  
 DR PROSITE: PS00631; CYTOSOL\_AP.1.  
 KM Hydrolyase: Aminopeptidase; Manganese: Complete proteome.  
 FT ACT\_SITE 210 210 POTENTIAL.  
 FT ACT\_SITE 284 284 POTENTIAL.  
 FT METAL 198 198 MANGANESE 2 (BY SIMILARITY).  
 FT METAL 203 203 MANGANESE 1 AND 2 (BY SIMILARITY).  
 FT METAL 221 221 MANGANESE 2 (BY SIMILARITY).  
 FT METAL 280 280 MANGANESE 1 (BY SIMILARITY).  
 FT METAL 282 282 MANGANESE 1 AND 2 (BY SIMILARITY).  
 SQ SEQUENCE 434 AA; 47285 MW; A5A592E51EB18B58 CRC64;  
 Query Match 73.7%; Score 28; DB 1; Length 434;  
 Best Local Similarity 66.7%; Pred. No. 1.9e+02;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

0Y 1 VLXDDLEA 9  
 DB 238 VSRDDLEA 246

0Y 1 VLXDDLEA 9  
 DB 286 VLADGLEA 294

```

CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M17.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AE006143; AAK03113.1; -.
CC DR HSSP: P00727; 1L4M.
CC DR MEROPS: M17.004; -.
CC DR InterPro: IPR000819; Peptidase_M17.
CC DR Pfam: PF00883; Peptidase_M17.1.
CC DR PRINTS: PR00481; LAMNOPPTDASE.
CC DR PROSITE: PS00631; CYTOSOL_AP.1.
CC KM Hydrolyase; aminopeptidase; Manganese; Complete proteome.
CC FT ACT_SITE 210 210 POTENTIAL.
CC FT METAL 198 198 MANGANESE 2 (BY SIMILARITY).
CC FT METAL 203 203 MANGANESE 1 AND 2 (BY SIMILARITY).
CC FT METAL 221 221 MANGANESE 2 (BY SIMILARITY).
CC FT METAL 280 280 MANGANESE 1 (BY SIMILARITY).
CC FT METAL 282 282 MANGANESE 1 AND 2 (BY SIMILARITY).
CC SQ SEQUENCE 434 AA; 46877 MW; 170A5FEA25AB9FB2 CRC64;

Query Match 73.7%; Score 28; DB 1; Length 434;
Best Local Similarity 66.7%; Pred. No. 1.9e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 LXDDLLEA 9
Db 286 VLADGLEIA 294

```

```

RESULT 44
MALH_FUSMR STANDARD; PRT; 441 AA.
ID MALH_FUSMR
AC 006901;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Maltose-6'-phosphate glucosidase (EC 3.2.1.122) (6-phospho-alpha-D-
DE glucosidase).
GN MALH.
OS Fusobacterium mortiferum.
OC Bacteria; Fusobacteria; Fusobacterium.
OX NCBI_TaxID=850;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 25557;
RA MEDLINE=97352666; PubMed=9209025;
RA Bouna C.L., Reizer J., Reizer A., Robrish S.A., Thompson J.;
RT "6-phospho-alpha-D-glucosidase from Fusobacterium mortiferum:
RT cloning, expression, and assignment to family 4 of the
RT glycosylhydrolases."
RL J. Bacteriol. 179:4129-4137(1997).
RN [2]
RP SEQUENCE OF 1-32, AND CHARACTERIZATION.
RC STRAIN=ATCC 25557;
RA MEDLINE=95247687; PubMed=7730284;
RA Thompson J., Gentry-Weeks C.R., Nguyen N.Y., Folk J.E., Robrish S.A.;
RT "Purification from Fusobacterium mortiferum ATCC 25557 of a 6-
RT phosphoryl-O-alpha-D-glucopyranosyl:6-phosphoglucosylhydrolase that
RT hydrolyzes maltose 6-phosphate and related phospho-alpha-D-
RT glucosides."
RL J. Bacteriol. 177:2505-2512(1995).
RN [3]
RP SUBSTRATE SPECIFICITY.
RC STRAIN=ATCC 25557;
RX PubMed=11882720;

```

```

RA Pikis A., Immel S., Robrish S.A., Thompson J.;
RT "Metabolism of sucrose and its five isomers by Fusobacterium
RT mortiferum."
RL Microbiology 148:843-852(2002).
CC -1- FUNCTION: Hydrolyzes a wide variety of 6-phospho-alpha-D-
CC glucosides including maltose-6'P, trehalose-6P and the 6'-
CC phosphorylated derivatives of the five linkage isomeric alpha-D-
CC glucosyl-D-fructoses: trehalulose-6'P, turanose-6'P, maltulose-
CC 6'P, leucrose-6'P, and palatinose-6'P. However, sucrose-6P is not
CC a substrate for malH, and this enzyme also fails to hydrolyze
CC beta-O-linked phosphorylated disaccharides such as cellobiose-6'P
CC and gentobiose-6'P.
CC -1- CATALYTIC ACTIVITY: Maltose 6'-phosphate + H(2)O = D-glucose 6-
CC phosphate + D-glucose.
CC -1- COFACTOR: NAD and a divalent metal ion. Manganese, iron, cobalt
CC and nickel ions enhance activity whereas magnesium, zinc, calcium
CC and strontium do not.
CC -1- PATHWAY: Sucrose isomers and related alpha-linked glucosides
CC metabolism.
CC -1- SUBUNIT: Homotetramer (By similarity).
CC -1- INDUCTION: By the five sucrose isomers and other alpha-glucosides
CC (but not by sucrose or glucose).
CC -1- MISCELLANEOUS: Optimum temperature is 40 degrees Celsius. Optimum
CC pH is 7.0-7.5.
CC -1- SIMILARITY: BELONGS TO FAMILY 4 OF GLYCOSYL HYDROLASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U81185; AAB63015.1; -.
CC DR InterPro: IPR001088; GH_4.
CC DR Pfam: PF02056; Glyco_hydro_4; 1.
CC DR PRINTS: PR00732; GLHYDRASE4.
CC DR PRODOM: PD006892; GH_4; 1.
CC DR PROSITE: PS01324; GLYCOSYL_HYDROL_F4; 1.
CC KM Hydrolyase; Glycosidase; NAD; Manganese; Iron; Cobalt; Nickel;
CC Carbohydrate metabolism.
CC FT NP_BIND 6 39 NAD (POTENTIAL);
CC FT ACT_SITE 39 39 PROTON DONOR (BY SIMILARITY).
CC FT ACT_SITE 109 109 NUCLEOPHILE (BY SIMILARITY).
CC FT ACT_SITE 358 358
CC SQ SEQUENCE 441 AA; 49712 MW; EE9DB5B35FA6AF46 CRC64;

Query Match 73.7%; Score 28; DB 1; Length 441;
Best Local Similarity 62.5%; Pred. No. 2e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 2 LXDDLLEA 9
Db 425 IJDDLLEA 432

```

```

RESULT 45
SYR_CORGL STANDARD; PRT; 550 AA.
ID SYR_CORGL
AC P35868; P41253;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Arginyl-tRNA synthetase (EC 6.1.1.19) (Arginine--tRNA ligase) (ARGRS).
GN ARRS OR CGL1179.
OS Corynebacterium glutamicum (Brevibacterium flavum).
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Corynebacteriaceae;
OC Corynebacterium.
OX NCBI_TaxID=1718;
RN [1]
RP SEQUENCE FROM N.A.

```

RC STRAIN-ATCC 13059 / AS019;  
 RX MEDLINE=91186817; PubMed=2082143;  
 RT Marcel T., Archer J.A.C., Mengin-Leceux D., Sinskey A.J.;  
 RT "Nucleotide sequence and organization of the upstream region of the  
 RT Corynebacterium glutamicum lysa gene.";  
 RL Mol. Microbiol. 4:1819-1830(1990).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-ATCC 13869;  
 RX MEDLINE=94042911; PubMed=8226683;  
 RA Ogura J.A., Malumbres M., Ertani G., Pisabarro A., Mateos L.M.,  
 RA Martin F., Martin J.F.;  
 RT "A gene encoding arginyl-tRNA synthetase is located in the upstream  
 RT region of the lysa gene in Brevibacterium lactofermentum: regulation  
 RT of argS-lysa cluster expression by arginine.";  
 RL J. Bacteriol. 175:7356-7362(1993).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-ATCC 13032 / DSM 20300 / NCIB 10025;  
 RX Nakagawa S.;  
 RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";  
 RT Submitted (Mar-2002) to the EMBL/Genbank/DBJ databases.  
 RL [4]  
 RP IDENTIFICATION.  
 RX MEDLINE=93268096; PubMed=8497194;  
 RA Sharp P.M., Mitchell K.J.;  
 RT "Corynebacterium glutamicum arginyl-tRNA synthetase.";  
 RL Mol. Microbiol. 8:200-200(1993).  
 CC -1- CATALYTIC ACTIVITY: ATP + L-arginine + tRNA(Arg) = AMP +  
 CC diphosphate + L-arginyl-tRNA(Arg).  
 CC -1- SUBUNIT: MONOMER (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: X54740; CAA38537.1; ALT\_INIT.  
 DR EMBL: 221501; CAA79710.1; -.  
 DR EMBL: AP005277; BAB98572.1; -.  
 DR PIR: S12227; S12227.  
 DR PIR: S42850; S42850.  
 DR InterPro: IPR001278; Arg-tRNA-synt\_1c.  
 DR InterPro: IPR005148; N.  
 DR InterPro: IPR001412; tRNA-synt\_1.  
 DR Pfam: PF00750; tRNA-synt\_1d; 1.  
 DR Pfam: PF03485; N-Arg\_1.  
 DR PRINTS: PRO1038; TRNASYNTHARG.  
 DR TIGRFS: TIGR00456; args; 1.  
 DR PROSITE: PS00178; AA\_TRNA\_LIGASE\_1; 1.  
 DR KMW: Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding.  
 FT SITE 130 140 "HIGH" REGION.  
 FT SITE 374 378 "KMSKS" REGION.  
 FT BINDING 377 377 ATP (BY SIMILARITY).  
 FT CONFLICT 355 355 G -> D (IN REF. 2).  
 FT CONFLICT 412 412 I -> M (IN REF. 2).  
 FT CONFLICT 513 513 V -> A (IN REF. 2).  
 FT CONFLICT 540 540 H -> R (IN REF. 2).  
 FT CONFLICT 540 540 H -> R (IN REF. 2).  
 SQ SEQUENCE 550 AA: 59723 MW: 3AF24BDEBDCAC1 CRC64;

Query Match 73.7%; Score 28; DB 1; Length 550;  
 Best Local Similarity 66.7%; Pred. No. 2.5e+02;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 VLXDDLEA 9  
 DB 383 VTLDLVEA 391

RESULT 46  
 MNT2\_YEAST  
 ID MNT2\_YEAST STANDARD; PRT; 558 AA.  
 AC P53059;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Alpha-1,3-mannosyltransferase MNT2 (EC 2.4.1.-).  
 GN MNT2 OR YGL257C OR NR0558.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 OX NCBITaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=S288c / FY1679;  
 RX MEDLINE=97127827; PubMed=8972578;  
 RA Colasac E., Maillier E., Robineau S., Netter P.;  
 RT "Sequence of a 39,411 bp DNA fragment covering the left end of  
 RT chromosome VII of Saccharomyces cerevisiae.";  
 RL Yeast 12:1555-1562(1996).  
 CC -1- FUNCTION: Mannosyltransferase involved in adding the 4th and 5th  
 CC mannose residues of O-linked glycans.  
 CC -1- PATHWAY: Glycosylation.  
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Golgi (Potential).  
 CC -1- SIMILARITY: BELONGS TO THE MNT1/MNT FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: X94357; CAA64130.1; -.  
 DR EMBL: Z72779; CAA96977.1; -.  
 DR SGD: S0003226; MNT2.  
 DR Transferrase: Glycosyltransferase; Glycoprotein; Transmembrane;  
 KW Signal-anchor; Golgi stack.  
 FT DOMAIN 1 6 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 7 27 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)  
 FT DOMAIN 1 6 (POTENTIAL).  
 FT DOMAIN 28 558 LUMENAL (POTENTIAL).  
 FT CARBOHYD 187 187 N-LINKED (GLCNAC... ) (POTENTIAL).  
 SQ SEQUENCE 558 AA: 64852 MW: 3E58ED62B4E29186 CRC64;

Query Match 73.7%; Score 28; DB 1; Length 558;  
 Best Local Similarity 55.6%; Pred. No. 2.5e+02;  
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 VLXDDLEA 9  
 DB 307 VASDDLES 315

RESULT 47

DHB4\_MOUSE

ID DHB4\_MOUSE STANDARD; PRT; 735 AA.

AC P51660;

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Estradiol 17 beta-dehydrogenase 4 (EC 1.1.1.62) (17-beta-HSD 4)

DE (17-beta-hydroxysteroid dehydrogenase 4).

GN HSD17B4 OR EDH17B4.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBITaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

```

RX MEDLINE=96135300; PubMed=8547180;
RA Normand T., Hussen B., Leenders F., Pelczar H., Baert J.-L.,
RA Beque A., Flourens A.C., Adamski J., de Lannoy Y.;
RT "Molecular characterization of mouse 17 beta-hydroxysteroid
RT dehydrogenase IV.";
RL J. Steroid Biochem. Mol. Biol. 55:541-548(1995).
CC -1- CATALYTIC ACTIVITY: Estradiol-17-beta + NAD(P)(+) = estrone +
CC NAD(P)H.
CC -1- PATHWAY: Testosterone biosynthesis; last step
CC -1- TISSUE SPECIFICITY: PRESENT IN MANY TISSUES WITH HIGHEST
CC CONCENTRATIONS IN LIVER AND KIDNEY.
CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
CC (SDR) FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X89998; CA62015.1; -
DR HSSP: 070351; 1E6W.
DR MGD: MGI:105089; Hsd17b4.
DR InterPro: IPR002198; ADH_short.
DR InterPro: IPR002539; MAOC_dehydratas.
DR InterPro: IPR003033; SCP2.
DR Pfam: PF00106; adh_short; 1.
DR Pfam: PF01575; MAOC_dehydratas; 1.
DR Pfam: PF02036; SCP2; 1.
DR PROSITE: PS00061; ADH_SHORT; 1.
DR Steroid biosynthesis; Oxidoreductase; NADP; Multigene family.
FW ACT_SITE 164 BY SIMILARITY.
FT SEQUENCE 735 AA; 79523 MW; E29AD3B6A268DA24 CRC64;
SQ
Query Match 73.7%; Score 28; DB 1; Length 735;
Best Local Similarity 66.7%; Pred. No. 3.4e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 VLXDDLLEA 9
Db 204 VLDPDLYEA 212
RESULT 48
YK69_YEAST STANDARD; PRT; 910 AA.
AC P36165;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypoetical 102.7 kDa protein in PRP16-SRP40 intergenic region.
GN YKR089C OR YKR409.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94262327; PubMed=8203164;
RA Garcia-Cantalejo J., Baladron V., Esteban P.F., Santos M.A., Bou G.,
RA Remacha M.A., Revuelta J.L., Baillesa J.P.G., Jimenez A., del Rey F.;
RT "The complete sequence of an 18,002 bp segment of Saccharomyces
RT cerevisiae chromosome XI contains the HBS1, MRP-L20 and PRP16 genes,
RT and six new open reading frames.";
RL Yeast 10:231-245(1994).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (potential).
CC -1- SIMILARITY: TO YEAST YMR313C AND S.POMBE SPC1450.16C.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z27116; CAAB1640.1; -
DR EMBL: Z28314; CAAB2168.1; -
DR PIR: S38167; S38167.
DR PIR: S39130; S39130.
DR SGD: S0001797; YKR089C.
DR InterPro: IPR002641; Patatin.
DR Pfam: PF01734; Patatin; 1.
KW Hypoetical protein; Transmembrane.
FT TRANSMEM 282 302 POTENTIAL.
FT TRANSMEM 310 326 POTENTIAL.
FT TRANSMEM 426 442 POTENTIAL.
FT SEQUENCE 910 AA; 102716 MW; 1CFC03CA46E4B9C CRC64;
SQ
Query Match 73.7%; Score 28; DB 1; Length 910;
Best Local Similarity 66.7%; Pred. No. 4.3e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 VLXDDLLEA 9
Db 686 VLDDNLLS 694
RESULT 49
OL56_STRAT STANDARD; PRT; 3519 AA.
AC Q07017;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Oleandomycin polyketide synthase, modules 5 and 6.
OS Streptomyces antibioticus.
OC Bacteria; Actinobacteria; Actinobacteriales; Streptomycetales; Streptomycetes.
OX NCBI_TaxID=1850;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94150470; PubMed=8107683;
RA Swan D.G., Rodriguez A.M., Vilches C., Mendez C., Salas J.A.;
RT "Characterisation of a Streptomyces antibioticus gene encoding a type
RT I polyketide synthase which has an unusual coding sequence.";
RL Mol. Gen. Genet. 242:358-362(1994).
CC -1- FUNCTION: MAY BE INVOLVED IN THE BIOSYNTHESIS OF THE OLEANDOMYCIN
CC LACTONE RING.
CC -1- COFACTOR: CONTAINS 2 COVALENTLY BOUND PHOSPHOPANTHEINES.
CC -1- SIMILARITY: CONTAINS 2 ACYL CARRIER DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L09654; AA19695.1; -
DR HSSP: P25715; 1MA.
DR InterPro: IPR001227; Ac transferase.
DR InterPro: IPR000794; Ketoacyl-synt.
DR InterPro: IPR003880; Pantine-attach.
DR InterPro: IPR001031; Thioesterase.
DR Pfam: PF00109; ketoacyl-synt; 2.
DR Pfam: PF00550; pp-binding; 2.
DR Pfam: PF00698; Acyl transf; 2.
DR Pfam: PF00975; Thioesterase; 1.
DR Pfam: PF02801; ketoacyl-synt_C; 2.
DR PROSITE: PS00012; PHOSPHOPANTHEINE; 2.
DR PROSITE: PS00060; B_KETOACYL_SYNTHASE; 2.

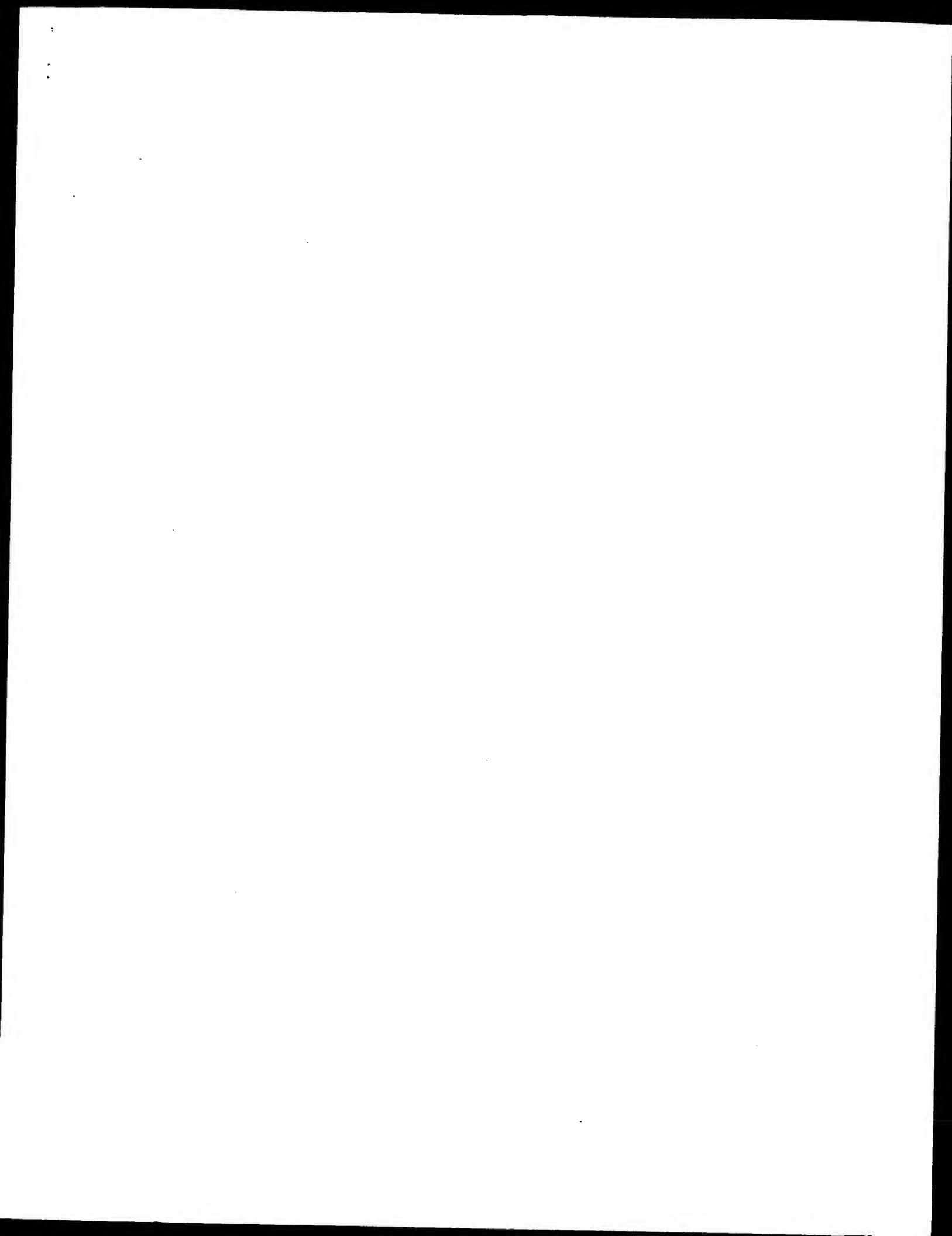
```

DR PROSITE: PS50075: ACP DOMAIN: 2.  
 KW Transferase: Acyltransferase; Antibiotic biosynthesis; NADP;  
 KM Phosphopantetheine; Multifunctional enzyme; Repeat.  
 FT DOMAIN 1 3519 ?  
 FT DOMAIN 32 501 ?  
 FT DOMAIN 569 891 ?  
 FT DOMAIN 1200 1382 ?  
 FT DOMAIN 1487 1561 ?  
 FT DOMAIN 1686 2156 ?  
 FT DOMAIN 2220 2541 ?  
 FT DOMAIN 2856 3038 ?  
 FT DOMAIN 3141 3215 ?  
 FT ACT\_SITE 3270 3519 ?  
 FT BINDING 210 210 ?  
 FT ACT\_SITE 660 660 ?  
 FT NP\_BIND 1203 1249 ?  
 FT BINDING 1524 1524 ?  
 FT ACT\_SITE 1859 1859 ?  
 FT ACT\_SITE 2311 2311 ?  
 FT NP\_BIND 2859 2905 ?  
 FT BINDING 3178 3178 ?  
 SQ SEQUENCE 3519 AA: 368561 MW: 41AE78AAEB61F86 CRC64;  
 Query Match 73.7%; Score 28; DB 1; Length 3519;  
 Best Local Similarity 75.0%; Pred. No. 1.8e+03;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 2 LXDDLLEA 9  
 DB 2421 LDELLEA 2428

RESULT 50  
 SACS\_HUMAN STANDARD; PRT: 3829 AA.  
 ID Q9NZJ4; O94835;  
 AC 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE SACS.  
 GN SACS OR KIA0730.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 OX NCBI\_TaxID-9606;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND VARIANT ALA-2619.  
 RA MEDLINE-20120709; PubMed-10655055;  
 RA Engert J.C., Berube P., Mercier J., Dore C., Lepage P., Ge B.,  
 RA Bouchard J.-P., Mathieu J., Melancon S.B., Schallin M., Lander E.S.,  
 RA Morgan K., Hudson T.J., Richter A.;  
 RA "ARSACS, a spastic ataxia common in northeastern Quebec, is caused by  
 RT mutations in a new gene encoding an 11.5-kb ORF."  
 RL Nat. Genet. 24.120-125(2000).  
 RN [2]  
 RP SEQUENCE OF 2826-3829 FROM N.A.  
 RC TISSUE-Brain;  
 RX MEDLINE-99087487; PubMed-9872452;  
 RA Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Miyajima N.,  
 RA Tanaka A., Kotani H., Nomura N., Ohara O.;  
 RT Prediction of the coding sequences of unidentified human genes. XI.  
 RT The complete sequences of 100 new cDNA clones from brain which code  
 RT for large proteins in vitro.;  
 RL DNA Res. 5:277-286(1998).  
 CC -1- FUNCTION: MAY FUNCTION IN CHAPERONE-MEDIATED PROTEIN FOLDING.  
 CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN THE CENTRAL NERVOUS  
 CC SYSTEM. ALSO FOUND IN SKELETAL MUSCLE AND AT LOW LEVELS IN  
 CC PANCREAS.  
 CC -1- DISEASE: DEFECTS IN SACS ARE THE CAUSE OF AUTOSOMAL RECESSIVE  
 CC SPASTIC ATAXIA OF CHARLEVOIX-SAGUENAY (ARSACS OR SACS). ARSACS IS

CC AN EARLY ONSET NEURODEGENERATIVE DISEASE WITH HIGH PREVALENCE IN  
 CC THE CHARLEVOIX-SAGUENAY-LAC-SAINT-JEAN REGION OF QUEBEC. IT IS  
 CC CHARACTERIZED BY ABSENT SENSORY-NERVE CONDUCTION, REDUCED MOTOR-  
 CC NERVE VELOCITY AND HYPERMYELINATION OF RETINAL-NERVE FIBERS.  
 CC -1- SIMILARITY: CONTAINS 1 J DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC at the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: AF193556; AAF31262.1; -;  
 CC EMBL: AB018273; BAA34450.1; -;  
 CC Genew: H0NC;10519; SACS.  
 CC MIM: 604490; -;  
 CC MIM: 270550; -;  
 CC InterPro: IPR001623; DnaJ\_N.  
 DR PROSITE: PS00636; DnaJ\_1; FALSE\_NEG.  
 DR PROSITE: PS50076; DnaJ\_2; 1.  
 KW Chaperone; Polymorphism.  
 FT DOMAIN 3556 3643  
 FT NP\_BIND 3556 3643  
 FT VARIANT 2619 2619  
 FT J-  
 FT V -> A.  
 FT /FTID-VAR\_010296.  
 SQ SEQUENCE 3829 AA: 436972 MW: 7AE990311E1E3E91 CRC64;  
 Query Match 73.7%; Score 28; DB 1; Length 3829;  
 Best Local Similarity 62.58%; Pred. No. 2e+03;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 VLXDDLLE 8  
 DB 1408 ILMDMLE 1415

Search completed: December 27, 2002, 14:43:21  
 Job time : 38 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 27, 2002, 14:40:34 : Search time 29 Seconds  
(without alignments)  
63.946 Million cell updates/sec

Title: US-09-489-760A-1  
Perfect score: 38  
Sequence: 1 VLXDDLFA 9

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 1000 summaries

Database :  
1: SP archaea: \*  
2: SP bacteria: \*  
3: SP fungi: \*  
4: SP human: \*  
5: SP invertebrate: \*  
6: SP mammal: \*  
7: SP mhc: \*  
8: SP organelle: \*  
9: SP phage: \*  
10: SP plant: \*  
11: SP rodent: \*  
12: SP virus: \*  
13: SP vertebrate: \*  
14: SP unclassified: \*  
15: SP virus: \*  
16: SP bacteriophage: \*  
17: SP archaea: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	36	94.7	1165	4	Q92619
2	34	89.5	616	17	Q27025
3	33	86.8	174	5	Q94256
4	32	84.2	112	2	Q31188
5	32	84.2	616	2	Q50407
6	32	84.2	797	3	Q9YTK6
7	32	84.2	894	10	Q9FWC7
8	32	84.2	4150	2	Q9KIVA
9	31	81.6	263	16	Q98L75
10	31	81.6	276	16	Q92A56
11	31	81.6	289	16	Q9KGF5
12	31	81.6	289	16	Q9KOG7
13	31	81.6	345	16	Q9JVG6
14	31	81.6	497	5	Q9ADK9
15	31	81.6	551	2	Q93G78
16	31	81.6	551	2	Q93G78

17	31	81.6	750	16	Q9KDP7	Q9Kd7 bacillus ha
18	31	81.6	1042	5	Q9UB20	Q9ub20 drosophila
19	31	81.6	1042	5	Q76911	Q76911 drosophila
20	31	81.6	1042	5	Q9W4W2	Q9w4w2 drosophila
21	30	78.9	72	5	Q9XND3	Q9xnd3 caenorhabdi
22	30	78.9	162	4	Q9H653	Q9h653 homo sapien
23	30	78.9	217	2	Q50241	Q50241 agrobacteri
24	30	78.9	248	10	Q9ASD4	Q9asd4 oryza sativ
25	30	78.9	253	5	Q9N3C0	Q9n3c0 caenorhabdi
26	30	78.9	302	2	Q9R6D2	Q9r6d2 agrobacteri
27	30	78.9	302	16	Q8U529	Q8u529 agrobacteri
28	30	78.9	327	16	Q93J73	Q93j73 streptomyce
29	30	78.9	361	10	Q9MI02	Q9mi02 arabidopsis
30	30	78.9	401	17	Q9H1H2	Q9h1h2 thermoplas
31	30	78.9	589	11	Q9E5E8	Q9e5e8 rattus norv
32	30	78.9	589	11	Q9QZC6	Q9qzc6 rattus norv
33	30	78.9	623	17	Q9V1X8	Q9v1x8 pyrococcus
34	30	78.9	633	16	Q9PD83	Q9pd83 xylella fas
35	30	78.9	799	10	Q9ACF8	Q9acf8 oryza sativ
36	30	78.9	962	12	Q89443	Q89443 african swi
37	30	78.9	963	12	Q8V9U2	Q8v9u2 african swi
38	30	78.9	1087	10	Q8SAX4	Q8sax4 oryza sativ
39	30	78.9	1201	16	Q8RG71	Q8rg71 fusobacteri
40	30	78.9	1410	2	Q9K1Z9	Q9k1z9 polyangium
41	30	78.9	1534	3	Q9P8H3	Q9p8h3 emericella
42	30	78.9	1999	16	Q8YWB9	Q8ywb9 anabaena sp
43	30	78.9	4976	2	Q87314	Q87314 mycobacteri
44	30	76.3	78	16	Q9TW27	Q9tw27 neisseria m
45	29	76.3	81	11	Q9JLB6	Q9jlb6 mus musculu
46	29	76.3	101	16	Q8UEB3	Q8ueb3 agrobacteri
47	29	76.3	103	2	Q53107	Q53107 mycoplasma
48	29	76.3	119	16	Q9X834	Q9x834 streptomyce
49	29	76.3	137	16	Q8ZR29	Q8zr29 salmonella
50	29	76.3	137	16	Q8ZK88	Q8zkr8 salmonella
51	29	76.3	137	16	Q8XBP9	Q8xbp9 escherichia
52	29	76.3	140	12	Q8QUV9	Q8quv9 african cas
53	29	76.3	145	10	Q94C10	Q94c10 arabidopsis
54	29	76.3	151	2	Q9EUX0	Q9eux0 haemophilus
55	29	76.3	151	2	Q9EY14	Q9ey14 actinobacill
56	29	76.3	156	5	Q18878	Q18878 caenorhabdi
57	29	76.3	178	16	Q9XOL1	Q9xol1 thermotoga
58	29	76.3	183	12	Q8V673	Q8v673 african cas
59	29	76.3	192	16	Q8YQ79	Q8yq79 anabaena sp
60	29	76.3	195	8	Q9T6L7	Q9t6l7 kluyveromyc
61	29	76.3	207	16	Q91588	Q91588 pseudomonas
62	29	76.3	212	8	Q21320	Q21320 saccharomyc
63	29	76.3	216	4	Q9C0J6	Q9c0j6 homo sapien
64	29	76.3	222	5	Q95XC3	Q95xc3 caenorhabdi
65	29	76.3	244	16	Q9CHC8	Q9chc8 lactococcus
66	29	76.3	251	8	Q9WK52	Q9wk52 candida gla
67	29	76.3	251	8	Q9GBE8	Q9gbe8 candida gla
68	29	76.3	251	8	Q8WBL2	Q8wbl2 candida gla
69	29	76.3	251	8	Q8WBL1	Q8wbl1 candida gla
70	29	76.3	251	8	Q8WBL0	Q8wbl0 candida gla
71	29	76.3	251	8	Q8WBL9	Q8wbl9 candida gla
72	29	76.3	251	8	Q8WBL8	Q8wbl8 candida gla
73	29	76.3	251	8	Q8WBL7	Q8wbl7 candida gla
74	29	76.3	265	12	Q36282	Q36282 pixuna viru
75	29	76.3	270	2	Q8VP47	Q8vp47 streptococ
76	29	76.3	277	17	Q26547	Q26547 methanobact
77	29	76.3	289	16	Q8XX79	Q8xx79 raietonia s
78	29	76.3	292	2	Q9PDS9	Q9pds9 streptococ
79	29	76.3	297	16	Q9A098	Q9a098 streptococ
80	29	76.3	321	16	Q8XT08	Q8xt08 raietonia s
81	29	76.3	323	16	Q8XC13	Q8xc13 escherichia
82	29	76.3	352	2	Q8VVP7	Q8vvp7 arcanobacte
83	29	76.3	358	10	Q8VZM0	Q8vzm0 arabidopsis
84	29	76.3	358	16	Q8YNS0	Q8yns0 anabaena sp
85	29	76.3	374	16	Q9CKE4	Q9cke4 pasteurella
86	29	76.3	385	17	Q8U453	Q8u453 pyrococcus
87	29	76.3	405	4	Q9H9T5	Q9h9t5 homo sapien
88	29	76.3	415	16	Q97NC8	Q97nc8 streptococ
89	29	76.3	429	5	Q9VT86	Q9vt86 drosophila

90	29	76.3	434	16	Q9KA09	Q9Kag9 bacillus ha	163	28	73.7	239	12	Q9WHY6	Q9wh6 influenza a
91	29	76.3	445	16	Q9J394	Q9j394 streptomyce	164	28	73.7	239	12	Q9WHY7	Q9wh7 influenza a
92	29	76.3	457	10	Q9MA13	Q9ma13 aradidopsis	165	28	73.7	239	12	Q9WHY8	Q9wh8 influenza a
93	29	76.3	465	16	Q93518	Q93518 salmonella	166	28	73.7	239	12	Q9WHY9	Q9wh9 influenza a
94	29	76.3	468	5	Q960C9	Q960c9 drosophila	167	28	73.7	239	12	Q9B0LW8	Q9b0l8 influenza a
95	29	76.3	477	16	Q92NL7	Q92nl7 rhizobium m	168	28	73.7	243	12	Q9B0LW8	Q9b0l8 influenza a
96	29	76.3	530	16	Q9KV80	Q9kv80 vibrio chol	169	28	73.7	245	12	Q9B0UJ7	Q9b0u7 influenza a
97	29	76.3	559	4	Q9H7V9	Q9h7v9 homo sapien	170	28	73.7	245	17	Q9H040	Q9h040 influenza a
98	29	76.3	559	4	Q96T11	Q96t11 homo sapien	171	28	73.7	246	12	Q9PZK6	Q9pzk6 influenza a
99	29	76.3	571	16	Q9A4J1	Q9a4j1 caulobacter	172	28	73.7	248	12	Q9B0LW3	Q9b0l3 influenza a
100	29	76.3	581	5	Q95R93	Q95r93 drosophila	173	28	73.7	248	12	Q9B0LW3	Q9b0l3 influenza a
101	29	76.3	584	4	Q9C0A9	Q9c0a9 homo sapien	174	28	73.7	250	12	Q9WHY2	Q9wh2 influenza a
102	29	76.3	587	16	Q8XU04	Q8xj04 clostridium	175	28	73.7	250	2	Q93P19	Q93p19 influenza a
103	29	76.3	602	10	Q9LW22	Q9lw22 oryza sativ	176	28	73.7	252	12	Q9BVL8	Q9bvl8 influenza a
104	29	76.3	602	16	Q8U966	Q8u966 agrobacteri	177	28	73.7	252	12	Q671A7	Q671a7 influenza a
105	29	76.3	617	2	Q87788	Q87788 pseudomonas	178	28	73.7	252	12	Q671A7	Q671a7 influenza a
106	29	76.3	619	11	Q62309	Q62309 mus musculu	179	28	73.7	252	12	Q671A7	Q671a7 influenza a
107	29	76.3	634	16	Q9EX29	Q9ex29 streptomyce	180	28	73.7	252	12	Q671A7	Q671a7 influenza a
108	29	76.3	638	16	Q92389	Q92389 streptomyce	181	28	73.7	252	12	Q671A7	Q671a7 influenza a
109	29	76.3	715	10	Q9SHS7	Q9shs7 aradidopsis	182	28	73.7	252	12	Q67204	Q67204 influenza a
110	29	76.3	722	10	Q9FJAO	Q9fjao aradidopsis	183	28	73.7	252	12	Q671B3	Q671b3 influenza a
111	29	76.3	747	5	Q9VDA4	Q9vda4 drosophila	184	28	73.7	252	12	Q671B3	Q671b3 influenza a
112	29	76.3	751	12	Q8OVFO	Q8ovf0 eyach virus	185	28	73.7	252	12	Q67202	Q67202 influenza a
113	29	76.3	809	16	Q8XZ24	Q8xz24 talstonia s	186	28	73.7	252	12	Q67175	Q67175 influenza a
114	29	76.3	846	10	Q9LR78	Q9lr78 aradidopsis	187	28	73.7	252	12	Q67175	Q67175 influenza a
115	29	76.3	897	10	Q9LW22	Q9lmw2 aradidopsis	188	28	73.7	252	12	Q67180	Q67180 influenza a
116	29	76.3	994	4	Q9UE60	Q9ue60 homo sapien	189	28	73.7	252	12	Q67180	Q67180 influenza a
117	29	76.3	1002	16	Q8YNE4	Q8yne4 anabaena sp	190	28	73.7	252	12	Q67171	Q67171 influenza a
118	29	76.3	1014	17	Q9HJ74	Q9hj74 thermoplasm	191	28	73.7	252	12	Q67332	Q67332 influenza a
119	29	76.3	1088	10	Q9SRB0	Q9srj0 aradidopsis	192	28	73.7	252	12	Q67200	Q67200 influenza a
120	29	76.3	1148	3	Q03897	Q03897 saccharomyc	193	28	73.7	252	12	Q67161	Q67161 influenza a
121	29	76.3	1174	2	Q932X6	Q932x6 lactococcus	194	28	73.7	252	12	Q67167	Q67167 influenza a
122	29	76.3	1253	12	Q9XK6	Q9xk6 venezuelan	195	28	73.7	252	12	Q67177	Q67177 influenza a
123	29	76.3	1330	16	Q8XR14	Q8xr14 talstonia s	196	28	73.7	252	12	Q67169	Q67169 influenza a
124	29	76.3	1408	8	Q8WT26	Q8wt26 psilotum nu	197	28	73.7	252	12	Q82711	Q82711 influenza a
125	29	76.3	1443	4	Q96AF0	Q96af0 homo sapien	198	28	73.7	252	12	Q9WM95	Q9wm95 influenza a
126	29	76.3	1498	4	Q99736	Q99736 homo sapien	199	28	73.7	252	12	Q9MB29	Q9mb29 influenza a
127	29	76.3	2412	4	Q92616	Q92616 homo sapien	200	28	73.7	252	12	Q9Q0M9	Q9q0m9 influenza a
128	29	76.3	2432	12	Q91QP4	Q91qp4 alchi virus	201	28	73.7	252	12	Q9Q0M7	Q9q0m7 influenza a
129	29	76.3	2433	12	Q91464	Q91464 alchi virus	202	28	73.7	252	12	Q9Q0M5	Q9q0m5 influenza a
130	29	76.3	3151	5	Q8SR52	Q8sr52 encephalito	203	28	73.7	252	12	Q997B0	Q997b0 influenza a
131	29	76.3	3151	5	Q8SR52	Q8sr52 encephalito	204	28	73.7	252	12	Q997A9	Q997a9 influenza a
132	28	73.7	104	17	Q9Y8L2	Q9y8l2 aeropyrum p	205	28	73.7	252	12	Q997A7	Q997a7 influenza a
133	28	73.7	112	2	Q9L8N0	Q9l8n0 haemophilus	206	28	73.7	252	12	Q997A6	Q997a6 influenza a
134	28	73.7	112	10	Q94125	Q94125 azotrizobiu	207	28	73.7	252	12	Q997A6	Q997a6 influenza a
135	28	73.7	122	10	Q24664	Q24664 porphyra pu	208	28	73.7	252	12	Q997A6	Q997a6 influenza a
136	28	73.7	135	12	Q65686	Q65686 beet necrot	209	28	73.7	252	12	Q997A6	Q997a6 influenza a
137	28	73.7	157	16	Q9KXRL	Q9kxrl bacillus ha	210	28	73.7	252	12	Q997A6	Q997a6 influenza a
138	28	73.7	163	16	Q9PPL6	Q9ppl6 campylobact	211	28	73.7	252	12	Q997A6	Q997a6 influenza a
139	28	73.7	164	5	Q8SUF7	Q8suf7 encephalito	212	28	73.7	252	12	Q997A6	Q997a6 influenza a
140	28	73.7	178	16	Q8YJX1	Q8yjx1 bruceella me	213	28	73.7	252	12	Q997A6	Q997a6 influenza a
141	28	73.7	180	16	Q98G25	Q98g25 rhizobium l	214	28	73.7	252	12	Q997A6	Q997a6 influenza a
142	28	73.7	180	16	Q8U9X1	Q8u9x1 agrobacteri	215	28	73.7	252	12	Q997A6	Q997a6 influenza a
143	28	73.7	191	16	Q92MG7	Q92mg7 rhizobium m	216	28	73.7	252	12	Q997A6	Q997a6 influenza a
144	28	73.7	195	3	Q60125	Q60125 schizosacch	217	28	73.7	252	12	Q997A6	Q997a6 influenza a
145	28	73.7	195	8	Q9TEK7	Q9tek7 kluveromyc	218	28	73.7	252	12	Q997A6	Q997a6 influenza a
146	28	73.7	198	12	Q8V6X4	Q8v6x4 influenza a	219	28	73.7	252	12	Q997A6	Q997a6 influenza a
147	28	73.7	199	16	Q92GG2	Q92gg2 influenza a	220	28	73.7	252	12	Q997A6	Q997a6 influenza a
148	28	73.7	209	16	Q9KRX5	Q9krx5 rickettsia	221	28	73.7	252	12	Q997A6	Q997a6 influenza a
149	28	73.7	213	16	Q9KRX5	Q9krx5 yersinia pe	222	28	73.7	252	12	Q997A6	Q997a6 influenza a
150	28	73.7	213	16	Q9KRX5	Q9krx5 yersinia pe	223	28	73.7	252	12	Q997A6	Q997a6 influenza a
151	28	73.7	220	12	Q93425	Q93425 salmonella	224	28	73.7	252	12	Q997A6	Q997a6 influenza a
152	28	73.7	220	12	Q93425	Q93425 salmonella	225	28	73.7	252	12	Q997A6	Q997a6 influenza a
153	28	73.7	228	12	Q8OLX8	Q8olx8 influenza a	226	28	73.7	252	12	Q997A6	Q997a6 influenza a
154	28	73.7	228	12	Q8OLX8	Q8olx8 influenza a	227	28	73.7	252	12	Q997A6	Q997a6 influenza a
155	28	73.7	228	12	Q8OLX8	Q8olx8 influenza a	228	28	73.7	252	12	Q997A6	Q997a6 influenza a
156	28	73.7	238	12	Q9WHY7	Q9wh7 influenza a	229	28	73.7	252	12	Q997A6	Q997a6 influenza a
157	28	73.7	239	12	Q9WHY7	Q9wh7 influenza a	230	28	73.7	252	12	Q997A6	Q997a6 influenza a
158	28	73.7	239	12	Q9WHY7	Q9wh7 influenza a	231	28	73.7	252	12	Q997A6	Q997a6 influenza a
159	28	73.7	239	12	Q9WHY7	Q9wh7 influenza a	232	28	73.7	252	12	Q997A6	Q997a6 influenza a
160	28	73.7	239	12	Q9WHY7	Q9wh7 influenza a	233	28	73.7	252	12	Q997A6	Q997a6 influenza a
161	28	73.7	239	12	Q9WHY7	Q9wh7 influenza a	234	28	73.7	252	12	Q997A6	Q997a6 influenza a
162	28	73.7	239	12	Q9WHY7	Q9wh7 influenza a	235	28	73.7	252	12	Q997A6	Q997a6 influenza a



236	28	73.7	252	12	Q9WBX8	Q9wbx8 influenza a	309	28	73.7	328	2	Q9AON7	Q9aon7 pseudomonas
237	28	73.7	252	12	Q9WBX0	Q9wbx0 influenza a	310	28	73.7	370	17	Q8TWI4	Q8twi4 methanopyru
238	28	73.7	252	12	Q9WBX2	Q9wbx2 influenza a	311	28	73.7	373	17	Q8TWI2	Q8twi2 methanocarc
239	28	73.7	252	12	Q9WBX4	Q9wbx4 influenza a	312	28	73.7	383	16	Q32505	Q32505 deinococcus
240	28	73.7	252	12	Q9WBX6	Q9wbx6 influenza a	313	28	73.7	389	5	Q9XTD7	Q9xtD7 caenorhabd
241	28	73.7	252	12	Q9WBX8	Q9wbx8 influenza a	314	28	73.7	397	10	Q9AVY3	Q9avY3 guillardia
242	28	73.7	252	12	Q9WBX2	Q9wbx2 influenza a	315	28	73.7	397	11	Q8ROC4	Q8roc4 mus musculu
243	28	73.7	252	12	Q9WBZ4	Q9wbz4 influenza a	316	28	73.7	398	16	Q8ZEP1	Q8zep1 salmonella
244	28	73.7	252	12	Q9WBZ6	Q9wbz6 influenza a	317	28	73.7	398	16	Q8ZEP9	Q8zep9 salmonella
245	28	73.7	252	12	Q9WBZ6	Q9wbz6 influenza a	318	28	73.7	398	16	Q8ZEP9	Q8zep9 streptomyc
246	28	73.7	252	12	Q9WBZ6	Q9wbz6 influenza a	319	28	73.7	401	16	Q9RDP8	Q9rdp8 streptomyc
247	28	73.7	252	12	Q9WBZ6	Q9wbz6 influenza a	320	28	73.7	407	16	Q9RDP5	Q9rdp5 streptomyc
248	28	73.7	252	12	Q9WBZ6	Q9wbz6 influenza a	321	28	73.7	411	10	Q9M310	Q9m310 arabisopsis
249	28	73.7	252	12	Q9WBZ6	Q9wbz6 influenza a	322	28	73.7	421	17	Q9TW60	Q9tw60 sulfolobus
250	28	73.7	252	12	Q9WBZ6	Q9wbz6 influenza a	323	28	73.7	426	11	Q9ERD5	Q9erd5 rattus norv
251	28	73.7	252	12	Q9WBZ6	Q9wbz6 influenza a	324	28	73.7	426	11	Q9ERD5	Q9erd5 rattus norv
252	28	73.7	252	12	Q9WBZ6	Q9wbz6 influenza a	325	28	73.7	439	4	Q9P2C0	Q9p2c0 clostridium
253	28	73.7	252	12	Q9WBZ6	Q9wbz6 influenza a	326	28	73.7	441	16	Q8XNT3	Q8xnt3 bacillus ha
254	28	73.7	252	12	Q9WBZ6	Q9wbz6 influenza a	327	28	73.7	442	16	Q9KGC1	Q9kgc1 streptomyc
255	28	73.7	252	12	Q9WBZ6	Q9wbz6 influenza a	328	28	73.7	460	16	Q9ACQ4	Q9acq4 streptomyc
256	28	73.7	252	12	Q9WBZ6	Q9wbz6 influenza a	329	28	73.7	486	16	Q9PPU0	Q9ppu0 streptomyc
257	28	73.7	252	12	Q9WBZ6	Q9wbz6 influenza a	330	28	73.7	522	16	Q9QXV7	Q9qxv7 staphylococ
258	28	73.7	252	12	Q9WBZ6	Q9wbz6 influenza a	331	28	73.7	523	11	Q9D5E4	Q9d5e4 mus musculu
259	28	73.7	252	12	Q9WBZ6	Q9wbz6 influenza a	332	28	73.7	524	2	Q8T901	Q8t901 chromatiu
260	28	73.7	252	12	Q9WBZ6	Q9wbz6 influenza a	333	28	73.7	538	16	P74340	P74340 synecocyst
261	28	73.7	252	12	Q9WBZ6	Q9wbz6 influenza a	334	28	73.7	543	17	Q9HNM1	Q9hnm1 halobacteri
262	28	73.7	252	12	Q9WBZ6	Q9wbz6 influenza a	335	28	73.7	556	3	Q9VWZ1	Q9vwz1 drosophila
263	28	73.7	252	12	Q9WBZ6	Q9wbz6 influenza a	336	28	73.7	563	5	Q9VWZ1	Q9vwz1 schizosacch
264	28	73.7	252	12	Q9WBZ6	Q9wbz6 influenza a	337	28	73.7	612	3	Q94560	Q94560 raltionia s
265	28	73.7	252	12	Q9WBZ6	Q9wbz6 influenza a	338	28	73.7	629	16	Q8XSP8	Q8xsp8 streptomyc
266	28	73.7	252	12	Q9WBZ6	Q9wbz6 influenza a	339	28	73.7	690	16	Q9LON2	Q9lon2 dermatophag
267	28	73.7	252	12	Q9WBZ6	Q9wbz6 influenza a	340	28	73.7	692	5	Q96720	Q96720 leishmania
268	28	73.7	252	12	Q9WBZ6	Q9wbz6 influenza a	341	28	73.7	700	11	Q9DBM3	Q9dbm3 mus musculu
269	28	73.7	252	12	Q9WBZ6	Q9wbz6 influenza a	342	28	73.7	735	11	Q9DBM3	Q9dbm3 mus musculu
270	28	73.7	252	12	Q9WBZ6	Q9wbz6 influenza a	343	28	73.7	769	16	Q8R955	Q8r955 thermoaer
271	28	73.7	252	12	Q9WBZ6	Q9wbz6 influenza a	344	28	73.7	776	16	Q8RLX5	Q8rlx5 listeria mo
272	28	73.7	252	12	Q9WBZ6	Q9wbz6 influenza a	345	28	73.7	776	16	Q9ZEB9	Q9zeb9 listeria in
273	28	73.7	252	12	Q9WBZ6	Q9wbz6 influenza a	346	28	73.7	776	16	Q8Y9E7	Q8y9e7 listeria mo
274	28	73.7	252	12	Q9WBZ6	Q9wbz6 influenza a	347	28	73.7	791	5	Q18325	Q18325 caenorhabd
275	28	73.7	252	12	Q9WBZ6	Q9wbz6 influenza a	348	28	73.7	837	4	Q9Y4R8	Q9y4r8 homo sapien
276	28	73.7	252	12	Q9WBZ6	Q9wbz6 influenza a	349	28	73.7	844	4	Q9Y168	Q9y168 homo sapien
277	28	73.7	252	12	Q9WBZ6	Q9wbz6 influenza a	350	28	73.7	867	16	Q9JYV4	Q9jyv4 neisseria m
278	28	73.7	252	12	Q9WBZ6	Q9wbz6 influenza a	351	28	73.7	867	16	Q9JYV4	Q9jyv4 neisseria m
279	28	73.7	252	12	Q9WBZ6	Q9wbz6 influenza a	352	28	73.7	889	10	Q9FHH9	Q9fhh9 arabidopsis
280	28	73.7	252	12	Q9WBZ6	Q9wbz6 influenza a	353	28	73.7	901	16	Q9GXP5	Q9gxp5 streptococc
281	28	73.7	252	12	Q9WBZ6	Q9wbz6 influenza a	354	28	73.7	980	5	Q96246	Q96246 plasmodium
282	28	73.7	252	12	Q9WBZ6	Q9wbz6 influenza a	355	28	73.7	997	3	Q9U0Y0	Q9u0y0 neospora
283	28	73.7	252	12	Q9WBZ6	Q9wbz6 influenza a	356	28	73.7	999	4	Q9HIS1	Q9his1 homo sapien
284	28	73.7	252	12	Q9WBZ6	Q9wbz6 influenza a	357	28	73.7	1021	4	Q94882	Q94882 homo sapien
285	28	73.7	252	12	Q9WBZ6	Q9wbz6 influenza a	358	28	73.7	1038	4	Q9HIS3	Q9his3 homo sapien
286	28	73.7	252	12	Q9WBZ6	Q9wbz6 influenza a	359	28	73.7	1092	4	Q9HIS4	Q9his4 homo sapien
287	28	73.7	252	12	Q9WBZ6	Q9wbz6 influenza a	360	28	73.7	1104	17	Q9VYX1	Q9vyx1 pyrococcus
288	28	73.7	252	12	Q9WBZ6	Q9wbz6 influenza a	361	28	73.7	1123	4	Q9UKX5	Q9ukx5 homo sapien
289	28	73.7	252	12	Q9WBZ6	Q9wbz6 influenza a	362	28	73.7	1123	4	Q9UKX5	Q9ukx5 homo sapien
290	28	73.7	252	12	Q9WBZ6	Q9wbz6 influenza a	363	28	73.7	1177	10	Q8S566	Q8s566 oryza sativ
291	28	73.7	252	12	Q9WBZ6	Q9wbz6 influenza a	364	28	73.7	1177	10	Q8S566	Q8s566 oryza sativ
292	28	73.7	252	12	Q9WBZ6	Q9wbz6 influenza a	365	28	73.7	1187	10	Q8S566	Q8s566 oryza sativ
293	28	73.7	252	12	Q9WBZ6	Q9wbz6 influenza a	366	28	73.7	1187	10	Q8S566	Q8s566 oryza sativ
294	28	73.7	252	12	Q9WBZ6	Q9wbz6 influenza a	367	28	73.7	1187	10	Q8S566	Q8s566 oryza sativ
295	28	73.7	252	12	Q9WBZ6	Q9wbz6 influenza a	368	28	73.7	1187	10	Q8S566	Q8s566 oryza sativ
296	28	73.7	252	12	Q9WBZ6	Q9wbz6 influenza a	369	28	73.7	1187	10	Q8S566	Q8s566 oryza sativ
297	28	73.7	252	12	Q9WBZ6	Q9wbz6 influenza a	370	28	73.7	1187	10	Q8S566	Q8s566 oryza sativ
298	28	73.7	252	12	Q9WBZ6	Q9wbz6 influenza a	371	28	73.7	1187	10	Q8S566	Q8s566 oryza sativ
299	28	73.7	252	12	Q9WBZ6	Q9wbz6 influenza a	372	28	73.7	1187	10	Q8S566	Q8s566 oryza sativ
300	28	73.7	252	12	Q9WBZ6	Q9wbz6 influenza a	373	28	73.7	1187	10	Q8S566	Q8s566 oryza sativ
301	28	73.7	252	12	Q9WBZ6	Q9wbz6 influenza a	374	28	73.7	1187	10	Q8S566	Q8s566 oryza sativ
302	28	73.7	252	12	Q9WBZ6	Q9wbz6 influenza a	375	28	73.7	1187	10	Q8S566	Q8s566 oryza sativ
303	28	73.7	252	12	Q9WBZ6	Q9wbz6 influenza a	376	28	73.7	1187	10	Q8S566	Q8s566 oryza sativ
304	28	73.7	252	12	Q9WBZ6	Q9wbz6 influenza a	377	28	73.7	1187	10	Q8S566	Q8s566 oryza sativ
305	28	73.7	252	12	Q9WBZ6	Q9wbz6 influenza a	378	28	73.7	1187	10	Q8S566	Q8s566 oryza sativ
306	28	73.7	252	12	Q9WBZ6	Q9wbz6 influenza a	379	28	73.7	1187	10	Q8S566	Q8s566 oryza sativ
307	28	73.7	252	12	Q9WBZ6	Q9wbz6 influenza a	380	28	73.7	1187	10	Q8S566	Q8s566 oryza sativ
308	28	73.7	252	12	Q9WBZ6	Q9wbz6 influenza a	381	28	73.7	1187	10	Q8S566	Q8s566 oryza sativ

382	28	73.7	3816	2	Q9KIV3	Q9KIV3 streptomyc	455	27	71.1	365	17	Q58321	Q58321 methanococ
383	27	71.1	48	10	Q9SWX0	Q9SWX0 salvia colu	456	27	71.1	368	10	Q8RU25	Q8RU25 oryza sativ
384	27	71.1	63	16	Q8XB66	Q8XB66 clostridium	457	27	71.1	368	17	Q8RU25	Q8RU25 zymomonas m
385	27	71.1	107	10	Q9SME0	Q9SME0 laminaria d	458	27	71.1	372	2	Q9RNP2	Q9RNP2 zymomonas m
386	27	71.1	108	6	Q62818	Q62818 ovis aries	459	27	71.1	373	17	Q8T2M0	Q8T2M0 pyrococcus
387	27	71.1	113	4	Q96EE4	Q96EE4 homo sapien	460	27	71.1	374	17	Q8T125	Q8T125 methanococ
388	27	71.1	117	2	Q9F719	Q9F719 chlorobium	461	27	71.1	377	8	Q09424	Q09424 chrysosplen
389	27	71.1	119	9	Q80057	Q80057 streptococ	462	27	71.1	379	17	Q27636	Q27636 methanobact
390	27	71.1	135	16	Q9RYK0	Q9RYK0 streptomyc	463	27	71.1	381	4	Q96K48	Q96K48 homo sapien
391	27	71.1	140	11	Q9D8D8	Q9D8D8 mus musculu	464	27	71.1	382	16	Q9KDE2	Q9KDE2 bacillus ha
392	27	71.1	151	11	P70607	P70607 rattus norv	465	27	71.1	384	12	Q65110	Q65110 adelaide ri
393	27	71.1	158	10	Q9LKJ9	Q9LKJ9 oryza sativ	466	27	71.1	385	16	Q9X201	Q9X201 thermotoga
394	27	71.1	158	10	Q40707	Q40707 oryza sativ	467	27	71.1	385	17	Q57906	Q57906 pyrococcus
395	27	71.1	161	5	Q90305	Q90305 anisakis si	468	27	71.1	385	17	Q57906	Q57906 chrysosplen
396	27	71.1	170	16	Q8UID1	Q8UID1 agrobacteri	469	27	71.1	392	8	Q36681	Q36681 quiniflora v
397	27	71.1	171	10	Q944C9	Q944C9 brassica ol	470	27	71.1	394	4	Q9H5S7	Q9H5S7 homo sapien
398	27	71.1	172	16	Q9UYA5	Q9UYA5 neisseria m	471	27	71.1	395	10	Q9LTR6	Q9LTR6 arabisdopsi
399	27	71.1	178	16	Q9RT89	Q9RT89 neisseria m	472	27	71.1	395	10	Q9X149	Q9X149 arabisdopsi
400	27	71.1	182	16	Q9ZSH6	Q9ZSH6 rhizobium m	473	27	71.1	396	8	Q09426	Q09426 chrysosplen
401	27	71.1	190	17	Q8U476	Q8U476 pyrococcus	474	27	71.1	409	16	Q97RH5	Q97RH5 chrysosplen
402	27	71.1	196	9	Q8SBX6	Q8SBX6 cyanophag	475	27	71.1	410	5	Q9VBA2	Q9VBA2 drosophila
403	27	71.1	196	9	Q8SBX5	Q8SBX5 cyanophag	476	27	71.1	410	5	Q9VBA2	Q9VBA2 drosophila
404	27	71.1	198	12	Q9YWM6	Q9YWM6 melanoplus	477	27	71.1	413	10	Q9M220	Q9M220 arabisdopsi
405	27	71.1	198	16	Q9Z7E5	Q9Z7E5 listeria in	478	27	71.1	413	11	Q9M401	Q9M401 arabisdopsi
406	27	71.1	198	16	Q8YF03	Q8YF03 brucella me	479	27	71.1	419	11	P97343	P97343 mus musculu
407	27	71.1	201	16	Q9ZC00	Q9ZC00 rickettsia	480	27	71.1	419	11	Q9CTT1	Q9CTT1 mus musculu
408	27	71.1	205	16	Q9A616	Q9A616 caulobacter	481	27	71.1	421	16	Q63285	Q63285 rattus norv
409	27	71.1	206	16	Q9W2X7	Q9W2X7 thermotoga	482	27	71.1	423	5	Q22310	Q22310 anabaena sp
410	27	71.1	213	17	Q59461	Q59461 pyrococcus	483	27	71.1	423	5	Q22310	Q22310 arabisdopsi
411	27	71.1	218	17	Q9VIR9	Q9VIR9 pyrococcus	484	27	71.1	429	10	Q93VY1	Q93VY1 arabisdopsi
412	27	71.1	221	16	Q8YGV5	Q8YGV5 brucella me	485	27	71.1	429	10	Q8W575	Q8W575 arabisdopsi
413	27	71.1	244	16	Q9K0D4	Q9K0D4 vibrio chol	486	27	71.1	432	8	Q8XSX1	Q8XSX1 escherichia
414	27	71.1	245	2	Q9E207	Q9E207 streptomyc	487	27	71.1	432	8	Q78605	Q78605 isocoria ver
415	27	71.1	248	16	Q9HXY9	Q9HXY9 pseudomonas	488	27	71.1	434	16	Q982M3	Q982M3 rhizobium l
416	27	71.1	252	2	Q8VVA6	Q8VVA6 corynebacte	489	27	71.1	435	16	Q97KG2	Q97KG2 clostridium
417	27	71.1	261	16	Q54172	Q54172 streptomyc	490	27	71.1	437	8	Q09416	Q09416 chrysosplen
418	27	71.1	262	16	Q9B8D8	Q9B8D8 candida alb	491	27	71.1	437	8	Q09422	Q09422 chrysosplen
419	27	71.1	263	16	Q97F03	Q97F03 streptococ	492	27	71.1	437	8	Q9BAM4	Q9BAM4 podocarpus
420	27	71.1	264	16	Q81249	Q81249 caenorhabdi	493	27	71.1	440	8	Q33148	Q33148 saxifraga p
421	27	71.1	275	16	Q8X014	Q8X014 escherichia	494	27	71.1	441	8	Q31822	Q31822 antrophium
422	27	71.1	279	10	Q9MIX2	Q9MIX2 arabisdopsi	495	27	71.1	445	3	Q96W31	Q96W31 candida alb
423	27	71.1	281	16	Q8ZAK7	Q8ZAK7 yersinia pe	496	27	71.1	445	16	Q9X1W4	Q9X1W4 thermotoga
424	27	71.1	283	2	Q05342	Q05342 yersinia ps	497	27	71.1	447	4	Q9ULJ4	Q9ULJ4 homo sapien
425	27	71.1	283	3	Q8WU66	Q8WU66 neurospora	498	27	71.1	449	16	Q99VSL	Q99VSL staphylococ
426	27	71.1	283	16	Q8UAV3	Q8UAV3 agrobacteri	499	27	71.1	454	2	P94251	P94251 borrelia bu
427	27	71.1	284	5	Q8SIO9	Q8SIO9 encephalito	500	27	71.1	454	16	Q50166	Q50166 borrelia bu
428	27	71.1	284	16	Q8R762	Q8R762 thermomane	501	27	71.1	455	10	Q9LTH4	Q9LTH4 arabisdopsi
429	27	71.1	288	16	Q8U783	Q8U783 agrobacteri	502	27	71.1	464	2	Q53886	Q53886 spirilloasma
430	27	71.1	295	2	Q9F5N4	Q9F5N4 burkholderi	503	27	71.1	465	8	Q31907	Q31907 bolandra or
431	27	71.1	296	9	Q80320	Q80320 bacterioph	504	27	71.1	465	8	Q32307	Q32307 elmera race
432	27	71.1	299	17	Q9V084	Q9V084 pyrococcus	505	27	71.1	465	8	Q32484	Q32484 jepsonia pa
433	27	71.1	310	16	Q97MF7	Q97MF7 clostridium	506	27	71.1	465	8	Q33136	Q33136 sullivanta
434	27	71.1	311	16	Q9RT05	Q9RT05 deinothococ	507	27	71.1	466	8	Q33136	Q33136 sullivanta
435	27	71.1	311	17	Q980M1	Q980M1 sulfolobus	508	27	71.1	467	8	Q20245	Q20245 achlys trip
436	27	71.1	313	12	Q8V305	Q8V305 swinepox vi	509	27	71.1	467	8	Q9TLN7	Q9TLN7 asterothrym
437	27	71.1	314	4	Q96C22	Q96C22 homo sapien	510	27	71.1	467	13	Q98TX3	Q98TX3 gallus gall
438	27	71.1	324	5	Q17551	Q17551 caenorhabdi	511	27	71.1	469	8	Q9CPD7	Q9CPD7 quinfina v
439	27	71.1	324	10	P93619	P93619 vigna ungui	512	27	71.1	469	8	Q95F21	Q95F21 broussaisia
440	27	71.1	324	16	Q9CEX1	Q9CEX1 lactococcus	513	27	71.1	469	8	Q95F19	Q95F19 platycrater
441	27	71.1	325	16	Q8X9K1	Q8X9K1 escherichia	514	27	71.1	472	10	Q94144	Q94144 oryza sativ
442	27	71.1	325	16	Q8UUK5	Q8UUK5 agrobacteri	515	27	71.1	476	8	Q9WTF0	Q9WTF0 tofieldia p
443	27	71.1	326	16	Q98N23	Q98N23 rhizobium l	516	27	71.1	476	8	Q93784	Q93784 humicola gr
444	27	71.1	326	16	Q8R197	Q8R197 fusobacteri	517	27	71.1	477	2	Q9Z474	Q9Z474 corynebacte
445	27	71.1	329	5	Q22060	Q22060 caenorhabdi	518	27	71.1	481	16	Q8Y5G3	Q8Y5G3 corynebacte
446	27	71.1	332	17	Q8Z260	Q8Z260 pyrococcus	519	27	71.1	485	16	Q8XMD5	Q8XMD5 oryza sativ
447	27	71.1	333	16	Q9XKZ3	Q9XKZ3 rhizobium m	520	27	71.1	486	10	Q94139	Q94139 oryza sativ
448	27	71.1	335	5	Q8S6T3	Q8S6T3 encephalito	521	27	71.1	488	16	Q8U931	Q8U931 agrobacteri
449	27	71.1	335	16	Q97R15	Q97R15 streptococ	522	27	71.1	496	9	Q9MBP0	Q9MBP0 staphylococ
450	27	71.1	338	10	Q94F27	Q94F27 arabisdopsi	523	27	71.1	496	9	Q8SDK3	Q8SDK3 staphylococ
451	27	71.1	340	5	Q8TBE4	Q8TBE4 trypanosoma	524	27	71.1	508	10	Q94145	Q94145 oryza sativ
452	27	71.1	341	9	Q9G082	Q9G082 bacterioph	525	27	71.1	513	9	Q9G083	Q9G083 bacterioph
453	27	71.1	344	4	Q8TAS1	Q8TAS1 homo sapien	526	27	71.1	515	16	Q92ND5	Q92ND5 rhizobium m
454	27	71.1	362	10	Q9FK08	Q9FK08 arabisdopsi	527	27	71.1	517	11	Q8R408	Q8R408 rattus norv

528	27	71.1	519	16	Q99XP6	Q99XP6 streptococ	601	27	71.1	5149	16	Q91179	Q91179 pseudomonas
529	27	71.1	520	2	Q9ZBA9	Q9ZBA9 ochrobactru	602	27	71.1	5825	10	Q82731	Q82731 vicia faba
530	27	71.1	524	16	Q98H34	Q98H34 rhizobium 1	603	26	68.4	25	12	Q91U61	Q91U61 influenza a
531	27	71.1	537	16	Q99Q24	Q99Q24 streptomyc	604	26	68.4	25	12	Q91U59	Q91U59 influenza a
532	27	71.1	545	10	Q64767	Q64767 arbidopsis	605	26	68.4	26	12	Q91U56	Q91U56 influenza a
533	27	71.1	548	4	Q8WV3	Q8WV3 homo sapien	606	26	68.4	34	12	Q91U63	Q91U63 influenza a
534	27	71.1	566	10	Q9C9H2	Q9C9H2 arbidopsis	607	26	68.4	41	12	Q91U54	Q91U54 influenza a
535	27	71.1	570	5	Q9XW62	Q9XW62 caenorhabd	608	26	68.4	43	12	Q91U58	Q91U58 influenza a
536	27	71.1	574	5	Q9VLS8	Q9VLS8 drosophila	609	26	68.4	43	12	Q91U52	Q91U52 influenza a
537	27	71.1	585	5	Q95TQ8	Q95TQ8 drosophila	610	26	68.4	54	12	Q8WUW2	Q8WUW2 influenza a
538	27	71.1	593	17	Q96YB8	Q96YB8 sulfolobus	611	26	68.4	63	4	Q9H4U5	Q9H4U5 homo sapien
539	27	71.1	596	10	Q8RWTS	Q8RWTS arbidopsis	612	26	68.4	73	12	Q57180	Q57180 vaccinia vi
540	27	71.1	601	5	Q9V6M8	Q9V6M8 drosophila	613	26	68.4	73	12	Q9JFE3	Q9JFE3 vaccinia vi
541	27	71.1	602	10	Q64582	Q64582 arbidopsis	614	26	68.4	73	12	Q8V534	Q8V534 monkeypox v
542	27	71.1	605	16	Q8ZH98	Q8ZH98 yersinia pe	615	26	68.4	73	12	Q8WUW4	Q8WUW4 cowpox viru
543	27	71.1	614	5	Q64149	Q64149 platyneris	616	26	68.4	73	16	Q97HU6	Q97HU6 clostridium
544	27	71.1	617	17	Q9YD4	Q9YD4 aeropyrum p	617	26	68.4	77	9	Q94M44	Q94M44 streptococ
545	27	71.1	626	16	Q8ZHN0	Q8ZHN0 yersinia pe	618	26	68.4	77	9	Q8W5Y5	Q8W5Y5 bacterioph
546	27	71.1	628	3	Q42939	Q42939 schizosacch	619	26	68.4	77	16	Q929A5	Q929A5 listeria in
547	27	71.1	652	12	Q8V3K5	Q8V3K5 swinepox vi	620	26	68.4	77	16	Q8Y507	Q8Y507 listeria mo
548	27	71.1	678	10	Q9C5H9	Q9C5H9 arbidopsis	621	26	68.4	82	16	Q8R4P3	Q8R4P3 mus musculu
549	27	71.1	683	10	Q8SAA8	Q8SAA8 sorghum bic	622	26	68.4	82	16	Q8YXZ8	Q8YXZ8 anabaena sp
550	27	71.1	683	10	Q9LFA3	Q9LFA3 arbidopsis	623	26	68.4	83	16	Q9A161	Q9A161 streptococ
551	27	71.1	685	16	Q9K100	Q9K100 neisseria m	624	26	68.4	87	10	Q8VX81	Q8VX81 pius pines
552	27	71.1	685	16	Q9JWP0	Q9JWP0 neisseria m	625	26	68.4	90	6	Q8WMR8	Q8WMR8 epiesicus f
553	27	71.1	686	6	Q9N0H4	Q9N0H4 sus scrofa	626	26	68.4	101	11	Q9CYO3	Q9CYO3 mus musculu
554	27	71.1	697	5	Q965W6	Q965W6 caenorhabd	627	26	68.4	106	16	Q82102	Q82102 salmonella
555	27	71.1	707	5	Q23191	Q23191 caenorhabd	628	26	68.4	118	17	Q27411	Q27411 methanobact
556	27	71.1	735	13	Q70529	Q70529 cavia porce	629	26	68.4	120	9	Q8SCW6	Q8SCW6 pseudomonas
557	27	71.1	737	13	Q8UVR8	Q8UVR8 oncorhynch	630	26	68.4	126	13	Q08515	Q08515 gallus gall
558	27	71.1	745	6	P79433	P79433 sus scrofa	631	26	68.4	126	16	Q9HE0	Q9HE0 rhizobium 1
559	27	71.1	756	10	Q8SAB1	Q8SAB1 sorghum bic	632	26	68.4	133	16	Q8Y129	Q8Y129 raltstonia s
560	27	71.1	832	16	Q98G05	Q98G05 rhizobium 1	633	26	68.4	138	16	Q8ZDY8	Q8ZDY8 yersinia pe
561	27	71.1	833	10	Q9LV01	Q9LV01 arbidopsis	634	26	68.4	139	17	Q8YX8	Q8YX8 methanopyru
562	27	71.1	840	11	Q9DC40	Q9DC40 mus musculu	635	26	68.4	144	2	Q8YVU6	Q8YVU6 rhodococcus
563	27	71.1	866	12	Q9WNG5	Q9WNG5 tobacco mos	636	26	68.4	146	16	Q9RX78	Q9RX78 delnocoous
564	27	71.1	866	12	Q9WNG5	Q9WNG5 tobacco mos	637	26	68.4	150	2	Q9RLX1	Q9RLX1 klebsiella
565	27	71.1	902	16	Q91742	Q91742 pseudomonas	638	26	68.4	151	2	Q9EUD1	Q9EUD1 haemophilus
566	27	71.1	978	16	Q9FNK9	Q9FNK9 campylobact	639	26	68.4	151	2	Q9EUC9	Q9EUC9 haemophilus
567	27	71.1	985	5	Q01590	Q01590 caenorhabd	640	26	68.4	151	2	Q9EUS0	Q9EUS0 haemophilus
568	27	71.1	1015	10	Q9AMR1	Q9AMR1 oryza sativ	641	26	68.4	151	2	Q9EUX8	Q9EUX8 haemophilus
569	27	71.1	1070	17	Q8TNY4	Q8TNY4 methanosarc	642	26	68.4	151	2	Q9EUX9	Q9EUX9 haemophilus
570	27	71.1	1091	16	Q9Z573	Q9Z573 streptomyc	643	26	68.4	151	2	Q9EUX8	Q9EUX8 haemophilus
571	27	71.1	1101	5	Q76369	Q76369 caenorhabd	644	26	68.4	151	2	Q9EUX8	Q9EUX8 haemophilus
572	27	71.1	1116	12	Q93A03	Q93A03 tobacco mos	645	26	68.4	151	2	Q9EUX8	Q9EUX8 haemophilus
573	27	71.1	1116	12	Q9QPN7	Q9QPN7 tobacco mos	646	26	68.4	151	2	Q9EUX8	Q9EUX8 haemophilus
574	27	71.1	1116	12	Q91BK4	Q91BK4 tobacco mos	647	26	68.4	151	2	Q9EUX8	Q9EUX8 haemophilus
575	27	71.1	1116	12	Q991T0	Q991T0 tomato mosa	648	26	68.4	151	2	Q9EUX8	Q9EUX8 haemophilus
576	27	71.1	1116	12	Q991B4	Q991B4 tobacco mos	649	26	68.4	151	2	Q9EUX8	Q9EUX8 haemophilus
577	27	71.1	1116	12	Q911B3	Q911B3 tobacco mos	650	26	68.4	151	2	Q9EUX8	Q9EUX8 haemophilus
578	27	71.1	1116	12	Q911B2	Q911B2 tobacco mos	651	26	68.4	151	2	Q9EUX8	Q9EUX8 haemophilus
579	27	71.1	1116	12	Q911B8	Q911B8 tomato mosa	652	26	68.4	151	2	Q9EUX8	Q9EUX8 haemophilus
580	27	71.1	1122	4	Q9HCG9	Q9HCG9 homo sapien	653	26	68.4	151	2	Q9EUX8	Q9EUX8 haemophilus
581	27	71.1	1168	5	Q09519	Q09519 caenorhabd	654	26	68.4	151	2	Q9EUX8	Q9EUX8 haemophilus
582	27	71.1	1206	16	Q9X1T0	Q9X1T0 thermotoga	655	26	68.4	151	2	Q9EUX8	Q9EUX8 haemophilus
583	27	71.1	1225	12	Q91IE0	Q91IE0 lymantria d	656	26	68.4	151	2	Q9EUX8	Q9EUX8 haemophilus
584	27	71.1	1225	12	Q993A4	Q993A4 bombyx mori	657	26	68.4	151	2	Q9EUX8	Q9EUX8 haemophilus
585	27	71.1	1346	5	Q9V6T8	Q9V6T8 drosophila	658	26	68.4	151	2	Q9EUX8	Q9EUX8 haemophilus
586	27	71.1	1464	12	Q8U0Z8	Q8U0Z8 rice black	659	26	68.4	151	2	Q9EUX8	Q9EUX8 haemophilus
587	27	71.1	1548	10	Q65531	Q65531 arbidopsis	660	26	68.4	151	2	Q9EUX8	Q9EUX8 haemophilus
588	27	71.1	1597	5	Q61346	Q61346 drosophila	661	26	68.4	151	2	Q9EUX8	Q9EUX8 haemophilus
589	27	71.1	1616	12	Q9JAO4	Q9JAO4 tobacco mos	662	26	68.4	151	2	Q9EUX8	Q9EUX8 haemophilus
590	27	71.1	1616	12	Q9J943	Q9J943 tobacco mos	663	26	68.4	151	2	Q9EUX8	Q9EUX8 haemophilus
591	27	71.1	1616	12	Q91BK5	Q91BK5 tobacco mos	664	26	68.4	151	2	Q9EUX8	Q9EUX8 haemophilus
592	27	71.1	1616	12	Q911B9	Q911B9 tomato mosa	665	26	68.4	151	2	Q9EUX8	Q9EUX8 haemophilus
593	27	71.1	1625	4	Q60303	Q60303 homo sapien	666	26	68.4	151	2	Q9EUX8	Q9EUX8 haemophilus
594	27	71.1	1722	5	Q19350	Q19350 caenorhabd	667	26	68.4	151	2	Q9EUX8	Q9EUX8 haemophilus
595	27	71.1	2136	10	Q8RYW8	Q8RYW8 oryza sativ	668	26	68.4	151	2	Q9EUX8	Q9EUX8 haemophilus
596	27	71.1	2197	12	Q91M09	Q91M09 a-2 plaque	669	26	68.4	151	2	Q9EUX8	Q9EUX8 haemophilus
597	27	71.1	2368	2	Q93TW6	Q93TW6 stigmatella	670	26	68.4	151	2	Q9EUX8	Q9EUX8 haemophilus
598	27	71.1	2712	10	Q9SB74	Q9SB74 arbidopsis	671	26	68.4	151	2	Q9EUX8	Q9EUX8 haemophilus
599	27	71.1	4621	11	Q8VHE6	Q8VHE6 mus musculu	672	26	68.4	151	2	Q9EUX8	Q9EUX8 haemophilus
600	27	71.1	4624	4	Q8TE73	Q8TE73 homo sapien	673	26	68.4	151	2	Q9EUX8	Q9EUX8 haemophilus

674	26	68.4	178	16	Q97N21	Q97n21 streptococc	747	26	68.4	252	12	Q96ON1	Q96on1 influenza a
675	26	68.4	185	16	Q9PAZ1	Q9paz1 xyella fas	748	26	68.4	252	12	Q9EAP0	Q9eap0 influenza a
676	26	68.4	185	16	Q97P65	Q97p65 streptococc	749	26	68.4	252	12	Q91B74	Q91b74 influenza a
677	26	68.4	188	6	Q9GME4	Q9gme4 callitrix	750	26	68.4	252	12	Q91U69	Q91u69 influenza a
678	26	68.4	191	4	Q9C062	Q9c062 homo sapien	751	26	68.4	252	12	Q91U68	Q91u68 influenza a
679	26	68.4	192	2	Q9Z5F6	Q9z5f6 morangella	752	26	68.4	252	12	Q91U67	Q91u67 influenza a
680	26	68.4	195	4	Q9H2J9	Q9h2j9 homo sapien	753	26	68.4	252	12	Q91U65	Q91u65 influenza a
681	26	68.4	195	11	Q9CZV8	Q9czv8 mus musculu	754	26	68.4	252	12	Q91CC2	Q91cc2 influenza a
682	26	68.4	196	9	Q8S8X4	Q8s8x4 mus musculu	755	26	68.4	252	12	Q91CC1	Q91cc1 influenza a
683	26	68.4	198	5	Q8S8U5	Q8s8u5 encephalito	756	26	68.4	252	12	Q91OH2	Q91oh2 influenza a
684	26	68.4	200	16	Q67107	Q67107 aquifex aeo	757	26	68.4	252	12	Q998K4	Q998k4 influenza a
685	26	68.4	200	17	Q8Z2Y0	Q8z2y0 pyrobaculum	758	26	68.4	252	12	Q91P48	Q91p48 influenza a
686	26	68.4	203	16	Q8YBS4	Q8ybs4 bruceella me	759	26	68.4	252	12	Q91P47	Q91p47 influenza a
687	26	68.4	204	11	Q9CYL9	Q9cyl9 mus musculu	760	26	68.4	252	12	Q91P48	Q91p48 influenza a
688	26	68.4	204	12	Q9MSW2	Q9msw2 tt vitus. o	761	26	68.4	252	12	Q9DIP3	Q9dip3 influenza a
689	26	68.4	210	10	Q9SSC7	Q9ssc7 arabidopsis	762	26	68.4	252	12	Q932J2	Q932j2 influenza a
690	26	68.4	210	16	Q9KAP8	Q9kap8 bacillus ha	763	26	68.4	253	2	Q93EY9	Q93ey9 streptococc
691	26	68.4	212	11	Q9JUB0	Q9jbu0 mus musculu	764	26	68.4	253	2	Q93EY9	Q93ey9 streptococc
692	26	68.4	213	2	Q68520	Q68520 myxococcus	765	26	68.4	253	5	Q95Y66	Q95y66 caenorhabdi
693	26	68.4	213	4	Q96763	Q96763 homo sapien	766	26	68.4	255	15	Q10258	Q10258 human immun
694	26	68.4	217	5	Q917D9	Q917d9 drosophila	767	26	68.4	259	2	Q91A42	Q91a42 streptococc
695	26	68.4	217	12	Q86801	Q86801 tomato aspe	768	26	68.4	259	2	Q93EY7	Q93ey7 streptococc
696	26	68.4	217	12	Q9YPR0	Q9ypr0 peanut stun	769	26	68.4	260	15	Q55328	Q55328 human immun
697	26	68.4	217	16	Q9YR09	Q9ypr9 peanut stun	770	26	68.4	260	15	Q56361	Q56361 human immun
698	26	68.4	217	16	Q9JN78	Q9jnt8 streptomyc	771	26	68.4	260	15	Q56364	Q56364 human immun
699	26	68.4	217	16	Q8Y1E8	Q8y1e8 raietonia s	772	26	68.4	261	16	Q8YOG8	Q8yog8 anabaena sp
700	26	68.4	218	12	Q86619	Q86619 tomato aspe	773	26	68.4	262	4	Q96AL9	Q96al9 homo sapien
701	26	68.4	218	12	Q8QPM3	Q8qpm3 apple mosai	774	26	68.4	262	5	Q96AL9	Q96al9 homo sapien
702	26	68.4	218	12	Q8QPM0	Q8qpm0 apple mosai	775	26	68.4	262	5	Q96AL9	Q96al9 homo sapien
703	26	68.4	219	1	Q8X248	Q8x248 haloarcula	776	26	68.4	264	17	Q92N13	Q92n13 rhizobium m
704	26	68.4	219	16	Q8YR80	Q8yrr80 bruceella me	777	26	68.4	264	17	Q92N13	Q92n13 rhizobium m
705	26	68.4	220	4	Q9NPD4	Q9npd4 homo sapien	778	26	68.4	265	16	Q91298	Q91298 pseudomonas
706	26	68.4	224	12	Q73450	Q73450 human papil	779	26	68.4	267	16	Q9K732	Q9k732 bacillus ha
707	26	68.4	226	4	Q96KS2	Q96ks2 homo sapien	780	26	68.4	268	4	Q96EW2	Q96ew2 homo sapien
708	26	68.4	226	16	Q9A3R9	Q9a3r9 caulobacter	781	26	68.4	268	5	Q9BUL6	Q9bul6 klebsiella
709	26	68.4	229	2	Q60108	Q60108 yersinia en	782	26	68.4	272	2	Q9Z5L0	Q9z5l0 klebsiella
710	26	68.4	230	2	Q9Y9J2	Q9y9j2 xenorhabdus	783	26	68.4	274	5	Q9W157	Q9w157 drosophila
711	26	68.4	230	12	Q9Y9Y9	Q9y9y9 melanoplus	784	26	68.4	282	12	Q9DS77	Q9ds77 beet necrot
712	26	68.4	231	16	Q8X1F4	Q8x1f4 salmonella	785	26	68.4	283	12	Q36303	Q36303 rice hoja b
713	26	68.4	233	2	Q9F5B4	Q9f5b4 agrobacteri	786	26	68.4	286	12	Q9ICE2	Q9ice2 rice hoja b
714	26	68.4	237	12	Q919Y9	Q919y9 influenza a	787	26	68.4	286	12	Q10390	Q10390 rice stripe
715	26	68.4	237	16	Q06978	Q06978 bacillus su	788	26	68.4	292	10	Q43752	Q43752 citrus sine
716	26	68.4	240	2	Q91700	Q91700 proteus mir	789	26	68.4	294	17	Q9M4T4	Q9m4t4 archaeoglob
717	26	68.4	240	16	Q8Z7F5	Q8z7f5 yersinia pe	790	26	68.4	296	10	Q912R5	Q912r5 arabidopsis
718	26	68.4	243	17	Q27016	Q27016 methanobact	791	26	68.4	297	11	Q8R5M0	Q8r5m0 mus musculu
719	26	68.4	244	12	Q9DRH7	Q9drh7 influenza a	792	26	68.4	298	2	Q9RME5	Q9rme5 zymomonas m
720	26	68.4	244	12	Q9DRH4	Q9drh4 influenza a	793	26	68.4	301	16	Q9AD04	Q9ad04 streptococc
721	26	68.4	244	12	Q9DRH8	Q9drh8 influenza a	794	26	68.4	304	16	Q9Z122	Q9z122 rhizobium m
722	26	68.4	244	12	Q9DRH5	Q9drh5 influenza a	795	26	68.4	304	16	Q8YJ16	Q8yji6 bruceella me
723	26	68.4	244	12	Q9DRH5	Q9drh5 influenza a	796	26	68.4	305	16	Q9RFG7	Q9rfg7 rhizobium l
724	26	68.4	244	12	Q9DRH3	Q9drh3 influenza a	797	26	68.4	307	17	Q9H6B0	Q9hp60 halobacteri
725	26	68.4	244	12	Q9DRH2	Q9drh2 influenza a	798	26	68.4	310	16	Q964W5	Q964w5 rhizobium l
726	26	68.4	244	12	Q9DRH1	Q9drh1 influenza a	799	26	68.4	312	4	Q8TF64	Q8tf64 homo sapien
727	26	68.4	244	12	Q9DRH0	Q9drh0 influenza a	800	26	68.4	312	12	Q65180	Q65180 african swi
728	26	68.4	244	12	Q9DRH0	Q9drh0 influenza a	801	26	68.4	312	16	Q9PA76	Q9paz76 xyella fas
729	26	68.4	244	12	Q9DRG9	Q9drg9 influenza a	802	26	68.4	316	16	Q8R753	Q8r753 thermotanaer
730	26	68.4	244	16	Q9BD84	Q9bd84 rhizobium l	803	26	68.4	319	10	Q9LEK9	Q9le66 arabidopsis
731	26	68.4	251	9	Q64332	Q64332 bacterioph	804	26	68.4	320	11	Q9CXA9	Q9cx9 mus musculu
732	26	68.4	251	12	Q9PZK0	Q9pzk0 influenza a	805	26	68.4	320	5	Q9N9Q4	Q9n9q4 leishmania
733	26	68.4	251	12	Q9PZK2	Q9pzk2 influenza a	806	26	68.4	320	11	Q9CXA9	Q9cx9 mus musculu
734	26	68.4	251	12	Q9PZK4	Q9pzk4 influenza a	807	26	68.4	320	13	Q9CXA9	Q9cx9 mus musculu
735	26	68.4	251	12	Q9PZK8	Q9pzk8 influenza a	808	26	68.4	323	2	Q52349	Q52349 escherichia
736	26	68.4	251	12	Q9PZL0	Q9pzl0 influenza a	809	26	68.4	323	2	Q93593	Q93593 xanthomonas
737	26	68.4	251	12	Q9PZL2	Q9pzl2 influenza a	810	26	68.4	324	9	Q8SPU5	Q8spu5 bacterioph
738	26	68.4	251	12	Q9PZL4	Q9pzl4 influenza a	811	26	68.4	324	9	Q8SPU5	Q8spu5 bacterioph
739	26	68.4	251	12	Q9PZL6	Q9pzl6 influenza a	812	26	68.4	324	9	Q940M5	Q940m5 arabidopsis
740	26	68.4	252	9	Q9MCS1	Q9mcs1 bacterioph	813	26	68.4	325	2	Q9L3P2	Q9l3p2 uncultured
741	26	68.4	252	12	Q9Q0P3	Q9q0p3 influenza a	814	26	68.4	325	16	Q99VQ1	Q99vq1 streptococc
742	26	68.4	252	12	Q9Q0P1	Q9q0p1 influenza a	815	26	68.4	326	5	Q76967	Q76967 podocoryne
743	26	68.4	252	12	Q9Q0N9	Q9q0n9 influenza a	816	26	68.4	327	10	Q9S536	Q9s536 arabidopsis
744	26	68.4	252	12	Q9Q0N7	Q9q0n7 influenza a	817	26	68.4	327	16	Q9CKE8	Q9cke8 pasteurella
745	26	68.4	252	12	Q9Q0N3	Q9q0n3 influenza a	818	26	68.4	327	16	Q8ZE46	Q8ze46 yersinia pe
746	26	68.4	252	12	Q9PZF0	Q9pzf0 influenza a	819	26	68.4	331	2	Q8YV05	Q8yv05 corynebacte

820	26	68.4	331	13	Q9DE30	Q9d30 xenopus lae	893	26	68.4	426	11	Q9E0T6	Q9e0t6 mus musculu
821	26	68.4	332	16	Q9E538	Q9e538 streptomyc	894	26	68.4	427	4	Q8WYR2	Q8wy2 homo sapien
822	26	68.4	333	17	Q8Z273	Q8z273 pyrobaculum	895	26	68.4	431	16	Q8U729	Q8u729 agrobacteri
823	26	68.4	334	17	Q8Z202	Q8z202 pyrobaculum	896	26	68.4	432	2	Q9RM63	Q9rm63 streptomyc
824	26	68.4	335	5	Q21935	Q21935 caenorhabdi	897	26	68.4	432	11	Q91KX5	Q91kx5 mus musculu
825	26	68.4	335	16	Q8X552	Q8x552 salmoneila	898	26	68.4	432	16	Q8UEC8	Q8uec8 agrobacteri
826	26	68.4	335	16	Q8X568	Q8x568 escherichia	899	26	68.4	434	3	Q9P456	Q9p456 aspergillus
827	26	68.4	335	16	Q8UBX7	Q8ubx7 agrobacteri	900	26	68.4	438	16	Q8Y9K6	Q8y9k6 listeria mo
828	26	68.4	335	17	Q8Z274	Q8z274 pyrobaculum	901	26	68.4	440	16	Q8XE10	Q8xe10 escherichia
829	26	68.4	336	13	Q9W6K9	Q9w6k9 odontaspis	903	26	68.4	446	16	Q9EBR5	Q9ebr5 homo sapien
830	26	68.4	336	13	Q9W6L1	Q9w6l1 ceterolius	904	26	68.4	448	10	Q40873	Q40873 picea glauc
831	26	68.4	336	17	Q8ZYD6	Q8zyd6 pyrobaculum	905	26	68.4	449	2	Q9AKI8	Q9aki8 rickettsia
832	26	68.4	336	17	Q8ZX14	Q8zx14 pyrobaculum	906	26	68.4	450	10	Q40844	Q40844 picea glauc
833	26	68.4	339	17	Q8ZTV7	Q8zTV7 pyrobaculum	907	26	68.4	450	10	Q40844	Q40844 picea glauc
834	26	68.4	341	3	P87171	P87171 schizosacch	908	26	68.4	452	4	Q43088	Q43088 schizosacch
835	26	68.4	341	10	Q91KA3	Q91ka3 arabidopsis	909	26	68.4	452	4	Q43088	Q43088 homo sapien
836	26	68.4	342	10	Q65004	Q65004 lycopersico	910	26	68.4	452	4	Q9HBA0	Q9hba0 homo sapien
837	26	68.4	344	17	Q980G7	Q980g7 sulfobolus	911	26	68.4	455	17	Q59191	Q59191 drosophila
838	26	68.4	345	2	Q48975	Q48975 mycoplasma	912	26	68.4	457	2	Q5ZHC3	Q5zhc3 streptococ
839	26	68.4	346	16	Q34788	Q34788 bacillus su	913	26	68.4	457	16	Q97PE6	Q97pe6 streptococ
840	26	68.4	347	4	Q75227	Q75227 homo sapien	914	26	68.4	457	16	Q92TF8	Q92tf8 rhizobium m
841	26	68.4	350	16	Q99ZM4	Q99zm4 streptococ	915	26	68.4	459	3	Q94729	Q94729 schizosacch
842	26	68.4	352	10	Q9FFH4	Q9ffh4 arabidopsis	916	26	68.4	459	3	Q94729	Q94729 clostridium
843	26	68.4	353	16	Q98JX0	Q98jx0 rhizobium 1	917	26	68.4	459	16	Q97126	Q97126 schizosacch
844	26	68.4	354	16	Q9RTJ4	Q9rtj4 deinococcus	918	26	68.4	460	5	Q96568	Q96568 caenorhabdi
845	26	68.4	355	10	Q9MT11	Q9mt11 arabidopsis	919	26	68.4	461	16	Q989F3	Q989f3 rhizobium 1
846	26	68.4	355	17	Q97AJ9	Q97aj9 thermoplas	920	26	68.4	463	5	Q9XTL0	Q9xtl0 caenorhabdi
847	26	68.4	356	10	Q9LPM9	Q9lpm9 arabidopsis	921	26	68.4	465	5	Q9Y0B9	Q9y0b9 caenorhabdi
848	26	68.4	357	10	Q9SVY8	Q9svy8 arabidopsis	922	26	68.4	466	2	Q938V7	Q938v7 bradyrhizob
849	26	68.4	362	17	Q9HJZ8	Q9hJz8 thermoplas	923	26	68.4	466	2	Q93785	Q93785 trichoderma
850	26	68.4	368	5	Q9XV95	Q9xv95 caenorhabdi	924	26	68.4	468	2	Q938U5	Q938u5 frankia sp.
851	26	68.4	371	16	Q8UIT9	Q8uit9 agrobacteri	925	26	68.4	468	2	Q938U5	Q938u5 neurospora
852	26	68.4	373	2	Q93716	Q93716 pantoea agg	926	26	68.4	470	2	Q93F84	Q93f84 pasteurella
853	26	68.4	373	17	Q9HNM3	Q9hnm3 halobacteri	927	26	68.4	472	2	Q93F84	Q93f84 pasteurella
854	26	68.4	374	10	Q9LUP9	Q9lup9 arabidopsis	928	26	68.4	472	2	Q93F84	Q93f84 pasteurella
855	26	68.4	377	2	Q93ST5	Q93st5 chlorobium	929	26	68.4	472	2	Q93F84	Q93f84 pasteurella
856	26	68.4	379	16	Q9RV61	Q9rv61 deinococcus	930	26	68.4	472	2	Q93F84	Q93f84 pasteurella
857	26	68.4	379	16	Q9RT82	Q9rt82 deinococcus	931	26	68.4	472	2	Q93F84	Q93f84 pasteurella
858	26	68.4	381	10	Q9ALV4	Q9alv4 oryza sativ	932	26	68.4	472	2	Q93F84	Q93f84 pasteurella
859	26	68.4	381	16	Q9K9F9	Q9k9f9 bacillus ha	933	26	68.4	472	2	Q93F84	Q93f84 pasteurella
860	26	68.4	387	11	Q9D344	Q9d344 mus musculu	934	26	68.4	472	16	Q9CLV0	Q9clv0 pasteurella
861	26	68.4	387	5	Q18546	Q18546 blomphalari	935	26	68.4	472	16	Q9CLV0	Q9clv0 pasteurella
862	26	68.4	387	5	Q8SVI6	Q8svi6 encephalito	936	26	68.4	480	5	Q9XU84	Q9xu84 caenorhabdi
863	26	68.4	387	16	Q8RRC3	Q8rrc3 fusobacteri	937	26	68.4	481	5	Q9UW55	Q9uw55 caenorhabdi
864	26	68.4	388	5	Q21174	Q21174 caenorhabdi	938	26	68.4	481	10	Q93ZK3	Q93zk3 drosophila
865	26	68.4	389	4	Q96AM4	Q96am4 homo sapien	939	26	68.4	481	10	Q93ZK3	Q93zk3 drosophila
866	26	68.4	389	16	Q86581	Q86581 streptomyc	940	26	68.4	482	10	Q94FND	Q94fnd lotus japon
867	26	68.4	389	16	Q9XMD8	Q9xmd8 streptomyc	941	26	68.4	484	16	Q9CKO2	Q9cko2 pasteurella
868	26	68.4	390	5	Q9XMD8	Q9xmd8 caenorhabdi	942	26	68.4	485	10	Q80401	Q80401 oryza sativ
869	26	68.4	393	5	P91093	P91093 caenorhabdi	943	26	68.4	488	10	Q94CC6	Q94cc6 arabidopsis
870	26	68.4	395	17	Q26656	Q26656 methanobact	944	26	68.4	493	10	Q9FMR2	Q9fmr2 arabidopsis
871	26	68.4	397	3	Q12123	Q12123 saccharomyc	945	26	68.4	493	17	Q28800	Q28800 archaeoglob
872	26	68.4	399	4	Q9B044	Q9b044 homo sapien	946	26	68.4	494	13	Q9W6L0	Q9w6l0 odontaspis
873	26	68.4	400	16	Q9A630	Q9a630 caulobacter	947	26	68.4	495	13	Q9W6L0	Q9w6l0 odontaspis
874	26	68.4	400	16	Q8UD01	Q8ud01 agrobacteri	948	26	68.4	495	17	Q9HS93	Q9hs93 halobacteri
875	26	68.4	402	4	Q9H3P0	Q9h3f0 homo sapien	949	26	68.4	496	13	Q9W6L3	Q9w6l3 glycine max
876	26	68.4	404	16	Q9LJA4	Q9lja4 streptomyc	950	26	68.4	496	13	Q9W6L3	Q9w6l3 lamia ditro
877	26	68.4	404	16	Q8XOJ3	Q8xoj3 ralsstonia s	951	26	68.4	497	13	Q9W6L3	Q9w6l3 atropias sup
878	26	68.4	406	10	Q8SA24	Q8sa24 oryza sativ	952	26	68.4	497	13	Q9W6L3	Q9w6l3 atropias pel
879	26	68.4	407	13	Q92032	Q92032 agkistrodon	953	26	68.4	498	10	Q9FHT5	Q9fht5 streptococ
880	26	68.4	408	10	Q9S628	Q9s628 arabidopsis	954	26	68.4	499	13	Q9W6L5	Q9w6l5 isurus oxyr
881	26	68.4	409	16	Q9LGH8	Q9lgh8 bruceella me	955	26	68.4	499	13	Q9W6L5	Q9w6l5 isurus oxyr
882	26	68.4	409	16	Q66367	Q66367 mycobacteri	956	26	68.4	499	13	Q9W6L5	Q9w6l5 isurus oxyr
883	26	68.4	411	16	Q9RC40	Q9rc40 bacillus ha	957	26	68.4	499	13	Q9W6L5	Q9w6l5 isurus oxyr
884	26	68.4	412	4	Q96N59	Q96n59 homo sapien	958	26	68.4	501	10	Q9ATN6	Q9atn6 anabaena sp
885	26	68.4	412	13	Q08514	Q08514 gallus gall	959	26	68.4	501	10	Q9ATN6	Q9atn6 anabaena sp
886	26	68.4	413	3	Q02783	Q02783 saccharomyc	960	26	68.4	503	16	Q8YML8	Q8yml8 streptomyc
887	26	68.4	417	16	Q910X9	Q910x9 pseudomonas	961	26	68.4	503	16	Q9K454	Q9k454 streptomyc
888	26	68.4	419	2	Q9LSW1	Q9lsw1 pterotella	962	26	68.4	509	16	Q9A2N4	Q9a2n4 caulobacter
889	26	68.4	421	17	Q9HIT9	Q9hit9 thermoplas	963	26	68.4	513	10	Q24245	Q24245 picea abies
890	26	68.4	422	10	Q9SHP0	Q9shp0 arabidopsis	964	26	68.4	513	16	Q9KSP9	Q9ksp9 vibrio chol
891	26	68.4	423	10	Q24742	Q24742 bacteroides	965	26	68.4	514	10	Q9ATL4	Q9atl4 castanea cr
892	26	68.4	423	10	Q8S9X8	Q8s9x8 oryza sativ							

```

966 26 68.4 514 10 094JTC2 094JTC2 oryza sativ
967 26 68.4 515 10 09CAB9 09CAB9 arabidopsis
968 26 68.4 525 16 09AFB8 09AFB8 arabidopsis
969 26 68.4 525 16 08UB13 08UB13 arabidopsis
970 26 68.4 529 17 09HLA5 09HLA5 human
971 26 68.4 534 13 09OXY5 09OXY5 human
972 26 68.4 540 2 09S3R0 09S3R0 human
973 26 68.4 546 12 066761 066761 human
974 26 68.4 546 12 089342 089342 human
975 26 68.4 546 12 091H63 091H63 human
976 26 68.4 546 12 091H63 091H63 human
977 26 68.4 546 12 091H63 091H63 human
978 26 68.4 546 12 091H63 091H63 human
979 26 68.4 546 12 091H63 091H63 human
980 26 68.4 546 12 091H63 091H63 human
981 26 68.4 546 12 091H63 091H63 human
982 26 68.4 546 12 091H63 091H63 human
983 26 68.4 546 12 091H63 091H63 human
984 26 68.4 546 12 091H63 091H63 human
985 26 68.4 546 12 091H63 091H63 human
986 26 68.4 546 12 091H63 091H63 human
987 26 68.4 546 12 091H63 091H63 human
988 26 68.4 546 12 091H63 091H63 human
989 26 68.4 546 12 091H63 091H63 human
990 26 68.4 546 12 091H63 091H63 human
991 26 68.4 546 12 091H63 091H63 human
992 26 68.4 546 12 091H63 091H63 human
993 26 68.4 546 12 091H63 091H63 human
994 26 68.4 546 12 091H63 091H63 human
995 26 68.4 546 12 091H63 091H63 human
996 26 68.4 546 12 091H63 091H63 human
997 26 68.4 546 12 091H63 091H63 human
998 26 68.4 546 12 091H63 091H63 human
999 26 68.4 546 12 091H63 091H63 human
1000 26 68.4 546 12 091H63 091H63 human

```

## ALIGNMENTS

```

RESULT 1
ID 092619 PRELIMINARY: PRT: 1165 AA.
AC 092619;
DT 01-FEB-1997 (TREMblrel. 02, Created)
DT 01-FEB-1997 (TREMblrel. 02, Last sequence update)
DE 01-MAR-2002 (TREMblrel. 20, Last annotation update)
DE MELOBLAST KIAA0223 (Fragment).
GN KIAA0223.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BONE MARROW; PubMed=9039502;
RA Nagase T., Seki N., Ishikawa K., Ohira M., Kawabayashi Y., Ohara O.,
RA Tanaka A., Kotani H., Miyajima N., Nomura N.;
RA "Prediction of the coding sequences of unidentified human genes. VI.
RT The coding sequences of 80 new genes (KIAA0201-KIAA0280) deduced by
RT analysis of cDNA clones from cell line KG-1 and brain.";
RL DNA Res. 3:321-329(1996).

```

```

RA Lamerdin J.E., McCreedy P.M., Skowronski E., Adamson A.W.,
RA Burkhardt-Schultz K., Gordon L., Kyle A., Ramirez M., Stillwagen S.,
RA Phan H., Velasco N., Ganes J., Dangnan L., Poundstone P.,
RA Christensen M., Georgescu A., Avila J., Liu S., Attix C., Andreise T.,
RA Trankheim M., Amico-Keller G., Coefield J., Duarte S., Lucas S.,
RA Bruce R., Thomas P., Quan G., Krommiller B., Arellano A.,
RA Montgomery M., Ow D., Nolan M., Trong S., Kobayashi A., Olsen A.O.,
RA Carrano A.V.;

```

```

RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: D86975; BAA1321.1; -.
DR EMBL: AC004151; AAC03237.1; -.
DR HSSP: 007960; IRCP.
DR InterPro: IPR001060; Cdc15_Fes_CIP4.
DR InterPro: IPR002219; DAG_PE-bind.
DR InterPro: IPR000198; RhogAP.
DR Pfam: PF00130; DAG_PE-bind; 1.
DR Pfam: PF00620; RhogAP; 1.
DR SMART: SM00109; CL; 1.
DR SMART: SM00055; FCH; 1.
DR SMART: SM00324; RhogAP; 1.
DR PROSITE: PS00479; DAG_PE_BIND_DOM_1; UNKNOWN_1.
DR PROSITE: PSS0081; DAG_PE_BIND_DOM_2; 1.
FT NON_TER
SQ SEQUENCE 1165 AA; 127344 MW; 92E7F6GCAFD458C9 CRC64;

```

```

Query Match 94.7%; Score 36; DB 4; Length 1165;
Best Local Similarity 88.9%; Pred. No. 64;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 VLXDDLEA 9
II |||||
Db 166 VLXDDLEA 174

```

```

RESULT 2
ID 027025 PRELIMINARY: PRT: 616 AA.
AC 027025;
DT 01-JUN-1998 (TREMblrel. 05, Created)
DT 01-JUN-1998 (TREMblrel. 05, Last sequence update)
DE 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Type I restriction modification enzyme, subunit M.
GN MTH942.
OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
OC Methanobacteriaceae; Methanothermobacter.
OX NCBI_Taxid=187420;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-DELTA H;
RA Smith D.R., Doucette-Stamm L.A., Delonghery C., Lee H.-M., Dubois J.,
RA Aldredge T., Bashirzaden R., Blakely D., Cook R., Gilbert K.,
RA Harrison D., Hoang L., Keagle P., Lamm W., Pochler B., Qiu D.,
RA Spadatore R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
RA Jiwani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
RA McDougall S., Shimer G., Goyal A., Pietrowski S., Church G.M.,
RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
RT "Complete genome sequence of Methanobacterium thermoautotrophicum
RT deltaH: functional analysis and comparative genomics.";
RL J. Bacteriol. 179:7135-7155(1997).

```

```

DR EMBL: AE000868; AAB85440.1; -.
DR InterPro: IPR003665; Methylase_M.
DR InterPro: IPR002296; N12N6_Mifase.
DR InterPro: IPR003356; N6_DNA_Mifase.
DR InterPro: IPR002052; N6_Mifase.
DR Pfam: PF02506; Methylase_M; 1.
DR Pfam: PF02384; N6_Mifase; 1.
DR PRINTS: PR00507; N12N6MIFASE.
DR PROSITE: PS00092; N6_MIFASE; UNKNOWN_1.
RW Complete proteome.
SQ SEQUENCE 616 AA; 71715 MW; 06DAFF1076A5D5BAD CRC64;

```

```

Query Match 89.5%; Score 34; DB 17; Length 616;
Best Local Similarity 77.8%; Pred. No. 86;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 VLXDDLEA 9
II |||||
Db 465 VLXDDLEA 473

```

## RESULT 3

Q94256 PRELIMINARY; PRT; 174 AA.  
 ID Q94256;  
 AC Q94256;  
 DT 01-FEB-1997 (TrEMBLrel. 02, Created)  
 DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE Hypothetical 19.2 kDa protein.  
 GN K04A8.9.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Pelodierinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RX MEDLINE=99069613; Pubmed=9851916;  
 RA None;  
 RT "Genome sequence of the nematode C. elegans: a platform for  
 investigating biology. The C. elegans Sequencing Consortium.";  
 RL Science 282:2012-2018(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA Geisel C., Bradshaw H.;  
 RT "The sequence of C. elegans cosmid K04A8.";  
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA Waterston R.;  
 RT "Direct Submission.";  
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: U64849; AAC4805.1; -;  
 DR InterPro: IPR000004; SAPB.  
 DR SMART: SM00118; SAPB; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 174 AA; 19210 MW; A6303CF6383BEBD9 CRC64;  
 Query Match 86.8%; Score 33; DB 5; Length 174;  
 Best Local Similarity 77.8%; Pred. No. 37;  
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 VLXDDLEA 9  
 II IIII I  
 Db 129 VLKDDLLDA 137

RESULT 4  
 O31188 PRELIMINARY; PRT; 112 AA.

ID O31188;  
 AC O31188;  
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE PII protein.  
 GN GLNB.  
 OS Rhodospirillum rubrum (Rhodospirillum rubrum).  
 OC Bacteria; Proteobacteria; alpha subdivision; Rhodospirillum group;  
 OC Rhodospirillum.  
 OX NCBI\_TaxID=1063;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-HR;  
 RX MEDLINE=98389683; Pubmed=9721307;  
 RA Qian Y., Tabita F.R.;  
 RT "Expression of glnb and a glnb-like gene (glnK) in a ribulose  
 biphosphate carboxylase/oxygenase-deficient mutant of Rhodospirillum  
 rubrum.";  
 RL EMBL: AF032116; AAC34721.1; -;  
 DR HSP; P05826; 2PIT.

DR InterPro: IPR002187; PII\_glnB.  
 DR InterPro: IPR002332; PII\_glnB\_UMP.  
 DR Pfam: PF00543; P-II; 1.  
 DR PRINTS: PR00340; PIIGLNB.  
 DR PRODOM: PD001194; PII\_glnB; 1.  
 DR PROSITE: PS00638; PII\_GLNBL\_CTER; 1.  
 DR PROSITE: PS00496; PII\_GLNBL\_UMP; 1.  
 SQ SEQUENCE 112 AA; 12142 MW; 85E0EENAC9EB13C4 CRC64;

Query Match 84.2%; Score 32; DB 2; Length 112;  
 Best Local Similarity 66.7%; Pred. No. 38;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

## RESULT 5

O50407 PRELIMINARY; PRT; 616 AA.  
 ID O50407;  
 AC O50407;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE ORF617 protein.  
 OS Mycobacterium smegmatis.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Corynebacteriaceae; Mycobacterium.  
 OX NCBI\_TaxID=1772;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-SN2;  
 RX MEDLINE=96118228; Pubmed=8574396;  
 RA Madhusudan K., Nagaraja V.;  
 RT "Mycobacterium smegmatis DNA gyrase: cloning and overexpression in  
 Escherichia coli.";  
 RL Microbiology 141:3029-3037(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-SN2;  
 RA Valakunja N.;  
 RL Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: X84077; CAA58886.1; -;  
 SQ SEQUENCE 616 AA; 64938 MW; 279AC173E042BABA CRC64;

Query Match 84.2%; Score 32; DB 2; Length 616;  
 Best Local Similarity 87.5%; Pred. No. 2.3e+02;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 VLXDDLE 8  
 II IIII I  
 Db 427 VLVDLLE 434

RESULT 6  
 O90TK6 PRELIMINARY; PRT; 797 AA.

ID O90TK6;  
 AC O90TK6;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE Hypothetical 91.7 kDa protein.  
 GN SPAC1486.03C.  
 OS Schizosaccharomyces pombe (fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomycetes.  
 OX NCBI\_TaxID=4896;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-972H-;  
 DR McDougall R.C., Rajandream M.A., Barrell B.G., Seeger K., Harris D.;



Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AL133357; CAB62413.1; -  
 DR InterPro: IPR000467; G\_patch.  
 DR Pfam: PF01585; G\_patch; 1.  
 DR SMART: SM00443; G\_patch; 1.  
 DR Hypothetical protein.  
 SO SEQUENCE 797 AA; 91669 MW; 6A54E06C35CD664A CRC64;

Query Match 84.2%; Score 32; DB 3; Length 797;  
 Best Local Similarity 77.8%; Pred. No. 3e+02;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 VLXDDLEA 9  
 DB 81 VTADDDLEA 89

## RESULT 7

Q9FWC7 PRELIMINARY; PRT; 894 AA.  
 AC Q9FWC7;  
 DT 01-MAR-2001 (TREMblrel. 16, Created)  
 DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)  
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)  
 DE Putative plant disease resistance polypeptide.  
 GN OSUNB50018B10.13.  
 OS Oryza sativa (Rice).  
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 CC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 CC Ehrhartoideae; Oryzaceae; Oryza.  
 NX NCBI\_TaxID=4530;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. NIPONBARE;  
 RA Buell J., Yuan Q., Moffat K.S., Hill J.N., Jenkins C.N., Burr P.C.,  
 RA Hsiao J., Zismann V., Pai G., Bowman C.L., Fujii C.Y., Vanaken S.E.,  
 RA Bowman C.L., Craven B., Utterback T.R., Khatai H., Feldblyum T.V.,  
 RA Quackenbush J., White O., Salzberg S.L., Fraser C.M.,  
 RT "Oryza sativa chromosome 10 BAC OSUNB0018B10 genomic sequence."  
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AC051634; AAG13423.1; -  
 DR InterPro: IPR005162; Retrotrans\_gag.  
 DR InterPro: IPR000477; RYase.  
 DR Pfam: PF03732; Retrotrans\_gag; 1.  
 DR Pfam: PF00078; rvt; 1.  
 DR KW Polyprotein: RNA-directed DNA polymerase.  
 SO SEQUENCE 894 AA; 100600 MW; 821DF61BE19B8B2E CRC64;

Query Match 84.2%; Score 32; DB 10; Length 894;  
 Best Local Similarity 66.7%; Pred. No. 3.4e+02;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 VLXDDLEA 9  
 DB 356 VLTDLLDA 364

## RESULT 8

Q9KIV4 PRELIMINARY; PRT; 4150 AA.  
 AC Q9KIV4;  
 DT 01-OCT-2000 (TREMblrel. 15, Created)  
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)  
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)  
 DE 8,8a-deoxyoleandolide synthase 1.  
 GN OLEAI.  
 OS Streptomyces antibioticus.  
 CC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 CC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.  
 NX NCBI\_TaxID=1890;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20363406; PubMed=10908114;

RA Shah S., Xue Q., Tang L., Carney J.R., Betlach M., McDaniel R.,  
 RT "Cloning, Characterization, and Heterologous Expression of a  
 RT Polyketide Synthase and P-450 Oxidase Involved in the Biosynthesis of  
 RT the Antibiotic Oleandomycin."  
 RL J. Antibiot. 53:502-508(2000).  
 DR EMBL: AF220951; AAF82408.1; -  
 DR HSSP: P25715; IMA

DR InterPro: IPR002106; MatRNA\_ligaseII.  
 DR InterPro: IPR001227; Ac\_transferase.  
 DR InterPro: IPR004410; Fabd.  
 DR InterPro: IPR000794; ketoacyl-synt.  
 DR InterPro: IPR003880; Ppanne\_attach.  
 DR InterPro: IPR000169; SHprot\_acsite.  
 DR Pfam: PF00698; Acyl\_transf. 3.  
 DR Pfam: PF00109; ketoacyl-synt. 3.  
 DR Pfam: PF02801; ketoacyl-synt.-C; 3.  
 DR Pfam: PF00550; pp-binding; 3.  
 DR TIGRfams: TIGR00128; fabd; 3.  
 DR PROSITE: PS00339; AA\_TRNA\_LIGASE\_II\_2; UNKNOWN\_1.  
 DR PROSITE: PS50075; ACP\_DOMAIN; 3.  
 DR PROSITE: PS00606; B\_KETOACYL-SYNTASE; 2.  
 DR PROSITE: PS00012; PHOSPHOPANTHETINE; 3.  
 DR PROSITE: PS00639; THIOL\_PROTEASE\_HIS; UNKNOWN\_1.  
 DR KW Phosphopantetheine; Transferase.  
 SO SEQUENCE 4150 AA; 435261 MW; 9383296C4C16647D CRC64;

Query Match 84.2%; Score 32; DB 2; Length 4150;  
 Best Local Similarity 87.5%; Pred. No. 1.7e+03;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 LXDLDLEA 9  
 DB 1781 LRDDLDLEA 1788

## RESULT 9

Q98L75 PRELIMINARY; PRT; 263 AA.  
 AC Q98L75;  
 DT 01-OCT-2001 (TREMblrel. 18, Created)  
 DT 01-OCT-2001 (TREMblrel. 18, Last sequence update)  
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)  
 DE ATP-binding protein of ABC transporter.  
 GN MLI149.  
 OS Rhizobium loti (Mesorhizobium loti).  
 CC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
 CC Phyllobacteriaceae; Mesorhizobium.  
 NX NCBI\_TaxID=381;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MAFF303099;  
 RX MEDLINE=21082930; PubMed=11214968;  
 RA Kaneko T., Nakamura Y., Sato S., Asanizu E., Kato T., Sasamoto S.,  
 RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,  
 RA Kishida Y., Kiyokawa C., Kohara M., Matsuno M., Matsuno A.,  
 RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpō S., Sugimoto M.,  
 RA Takeuchi C., Yamada M., Tabata S.,  
 RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium  
 RT Mesorhizobium loti."  
 RL DNA Res. 7:331-338(2000).  
 DR EMBL: AP002996; BAB48588.1; -  
 DR InterPro: IPR003593; AAA\_ATPase.  
 DR InterPro: IPR003439; AAA\_transport.  
 DR Pfam: PF00005; ABC\_tran; 1.  
 DR Prodom: PD000006; ABC\_transport; 1.  
 DR SMART: SM00382; AAA; 1.  
 DR PROSITE: PS00211; ABC\_TRANSPORTER; 1.  
 DR ATP-binding; Complete proteome.  
 KW SEQUENCE 263 AA; 28452 MW; 42B3032FF840DE9E CRC64;

Query Match 81.6%; Score 31; DB 16; Length 263;  
 Best Local Similarity 75.0%; Pred. No. 1.5e+02;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;



OY 1 VLXDDLLE 8  
111111  
Db 229 ISDDLLE 236

## RESULT 10

Q92AS6 PRELIMINARY: PRT: 276 AA.  
ID 092AS6  
AC 092AS6  
DT 01-DEC-2001 (TREMblrel. 19, last sequence update)  
DT 01-DEC-2001 (TREMblrel. 19, last sequence update)  
DE 01-MAR-2002 (TREMblrel. 20, last annotation update)  
DE Hypothetical protein lin1843.  
GN LIN1843.  
OS Listeria innocua.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;  
OC Listeriaceae; Listeria.  
OX NCBI\_TaxID=1642;  
[1]

SEQUENCE FROM N.A.  
RC STRAIN-CLIP 11262 / SEROVAR 6A;  
RX PubMed=11679669;  
RA Glaser P., Frangoul L., Buchrieser C., Rusniok C., Amend A.,  
Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,  
Charbit A., Chetouani F., Couve E., de Daruvar A., Denoux P.,  
Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,  
Entian K.-D., Fahl H., Garcia-del Portillo F., Garrido P.,  
Gautier L., Goebel W., Gomez-Lopez N., Hain T., Haut J., Jackson D.,  
Jones L.-M., Kaerst U., Kreft J., Kunz M., Kunst F., Kurapkhat G.,  
Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,  
Nordisiek G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,  
Rommel B., Rose M., Schluter T., Simoes N., Tierrez A.,  
Vazquez-Boland J.-A., Voss H., Wehlend J., Cossart P.;  
"Comparative genomics of Listeria species.";  
RT Science 294:849-852(2001).  
RL EMBL: AL591981; CAC99810.1; -  
DR Listeria; LIN1843; -  
DR InterPro: IPR000515; BPD\_transp.  
DR Pfam: PF00528; BPD\_transp; 1.  
DR PROSITE: PS00402; BPD\_TRANSF\_INN\_MEMBER; UNKNOWN\_1.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 276 AA; 31102 MW; 3BDE99E47DB9B2FA CRC64;

Query Match 81.6%; Score 31; DB 16; Length 276;  
Best Local Similarity 75.0%; Pred. No. 1.6e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 LXDLLLEA 9  
111111  
Db 165 ISDDLLEA 172

## RESULT 11

Q8Y6F5 PRELIMINARY: PRT: 276 AA.  
ID 08Y6F5  
AC 08Y6F5  
DT 01-MAR-2002 (TREMblrel. 20, last sequence update)  
DT 01-MAR-2002 (TREMblrel. 20, last sequence update)  
DE 01-JUN-2002 (TREMblrel. 21, last annotation update)  
DE Hypothetical protein lmo1732.  
GN LMO1732.  
OS Listeria monocytogenes.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;  
OC Listeriaceae; Listeria.  
OX NCBI\_TaxID=1639;  
[1]

SEQUENCE FROM N.A.  
RC STRAIN-EGD-E / SEROVAR 1/2A;  
RX MEDLINE=21537273; PubMed=11679669;  
RA Glaser P., Frangoul L., Buchrieser C., Rusniok C., Amend A.,  
Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,  
Charbit A., Chetouani F., Couve E., de Daruvar A., Denoux P.,

Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,  
Entian K.-D., Fahl H., Garcia-del Portillo F., Garrido P.,  
Gautier L., Goebel W., Gomez-Lopez N., Hain T., Haut J., Jackson D.,  
Jones L.-M., Kaerst U., Kreft J., Kunz M., Kunst F., Kurapkhat G.,  
Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,  
Nordisiek G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,  
Rommel B., Rose M., Schluter T., Simoes N., Tierrez A.,  
Vazquez-Boland J.-A., Voss H., Wehlend J., Cossart P.;  
"Comparative genomics of Listeria species.";  
RT Science 294:849-852(2001).  
RL EMBL: AL591981; CAC99810.1; -  
DR Listeria; LMO1732; -  
DR InterPro: IPR000515; BPD\_transp.  
DR Pfam: PF00528; BPD\_transp; 1.  
DR PROSITE: PS00402; BPD\_TRANSF\_INN\_MEMBER; UNKNOWN\_1.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 276 AA; 31088 MW; 808E99E56DE9B2F8 CRC64;

Query Match 81.6%; Score 31; DB 16; Length 276;  
Best Local Similarity 75.0%; Pred. No. 1.6e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 LXDLLLEA 9  
111111  
Db 165 ISDDLLEA 172

## RESULT 12

Q9K0G7 PRELIMINARY: PRT: 289 AA.  
ID 09K0G7  
AC 09K0G7  
DT 01-OCT-2000 (TREMblrel. 15, last sequence update)  
DT 01-OCT-2000 (TREMblrel. 15, last sequence update)  
DE 01-JUN-2002 (TREMblrel. 21, last annotation update)  
DE UTP-glucose-1-phosphate uridylyltransferase.  
GN NMB0638.  
OS Neisseria meningitidis (serogroup B).  
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.  
OX NCBI\_TaxID=491;  
[1]

SEQUENCE FROM N.A.  
RC STRAIN-MC58 / SEROGROUP B;  
RX MEDLINE=20175755; PubMed=10710307;  
RA Tettelein H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,  
Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,  
Nelson W.C., Gwin M.L., DeBoy R., Peterson J.D., Hickey E.K.,  
Hait D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,  
Mason T., Ciecko A., Parksey D.S., Blair E., Clifton H., Clark E.B.,  
Gill J., Scariato V., Maignani V., Pizzi M., Grandi G., Sun L.,  
Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;  
"Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.";  
RT Science 287:1809-1815(2000).  
RL EMBL: AE002419; AAF41061.1; -  
DR TIGR: NMB0638; -  
DR InterPro: IPR001825; NTP\_transferase.  
DR Pfam: PF00483; NTP\_transferase; 1.  
DR TIGRFAMs: TIGR01099; galu; 1.  
KW Transferase; Nucleotidyltransferase; Complete proteome.  
SQ SEQUENCE 289 AA; 32030 MW; D8D76E2C692781FF CRC64;

Query Match 81.6%; Score 31; DB 16; Length 289;  
Best Local Similarity 55.6%; Pred. No. 1.7e+02;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 VLXDDLLEA 9  
111111  
Db 130 ILADLLIDA 138

## RESULT 13

Q9JVG6

```

ID 09JVG6 PRELIMINARY; PRT; 289 AA.
AC 09JVG6;
DT 01-OCT-2000 (TREMUREL. 15, Created)
DT 01-OCT-2000 (TREMUREL. 15, Last sequence update)
DT 01-JUN-2002 (TREMUREL. 21, Last annotation update)
DE Putative UDP-glucose-1-phosphate uridylyltransferase.
GN CAU OR NM00848.
OS Neisseria meningitidis (serogroup A).
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
NCBI_TaxID=65699;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=22491 / SEROGROUP A / SEROTYPE 4A;
RX MEDLINE=20222556; PubMed=10761919;
RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
RA Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holroyd S.,
RA Jagels K., Leather S., Moule S., Mungall K., Quail M.A.,
RA Rajandream M.A., Rutherford K.M., Simmonds M., Skellton J.,
RA Whitehead S., Spratt B.G., Barrell B.G.;
RT "Complete DNA sequence of a serogroup A strain of Neisseria
RT meningitidis 22491."
RL Nature 404:502-506(2000).
DR EMBL: AL162754; CAB84129.1;
DR InterPro: IPR001825; NTP transferase.
DR Pfam: PF00483; NTP transferase; 1.
DR TIGRPFAMs: TIGR01099; galU; 1.
KW Transference, Nucleotidyltransferase; Complete proteome.
SQ SEQUENCE 289 AA; 32102 MW; ACCAEC2C810D8D08 CRC64;

Query Match
Best Local Similarity 81.6%; Score 31; DB 16; Length 289;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLXDDLLEA 9
DB 130 ILADLDDA 138

RESULT 14
Q9ADK9 PRELIMINARY; PRT; 345 AA.
AC 09ADK9;
DT 01-JUN-2001 (TREMUREL. 17, Created)
DT 01-JUN-2001 (TREMUREL. 17, Last sequence update)
DT 01-JUN-2002 (TREMUREL. 21, Last annotation update)
DE Putative aldoketoreductase.
GN 2SCK31.11C OR SC04951.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Oliver K., Harris D.;
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Cerdano A.M., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RX MEDLINE=97000351; PubMed=8943436;
RA Redenbach M., Kieser H.M., Denapate D., Eichner A., Cullum J.,
RA Kinschi H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome."
RL Mol. Microbiol. 21:77-96(1996).
RN [4]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN=A3(2) / ML45;
RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2)."
RL Nature 417:141-147(2002).
DR EMBL: AL451182; CAC18692.2;
DR EMBL: AL512657; CAC30937.1;
DR HSSP: P06632; 1HW6.
DR InterPro: IPR001395; Aldo/ket_red.
DR Pfam: PF00248; aldo_ket_red; 1.
DR PRINTS: PR00069; ALDKETRDASE.
DR PRINTS: PR01577; KCNABCHANNEL.
DR ProDom: PD000288; Aldo/ket_red; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 345 AA; 37200 MW; 1702711E71DF697E CRC64;

Query Match
Best Local Similarity 81.6%; Score 31; DB 16; Length 345;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLXDDLLEA 9
DB 300 VLSDVDLDA 308

RESULT 15
O01964 PRELIMINARY; PRT; 497 AA.
AC 001964;
DT 01-JUL-1997 (TREMUREL. 04, Created)
DT 01-JUL-1997 (TREMUREL. 04, Last sequence update)
DT 01-JUN-2002 (TREMUREL. 21, Last annotation update)
DE Hypothetical 57.3 kDa protein.
GN C27A12.7.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium."
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Murray J., Wohlmann P., Biewald T.;
RT "The sequence of C. elegans cosmid C27A12."
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Waterston R.;
RT "Direct Submission."
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF003137; AAB93644.1;
DR InterPro: IPR002867; Znf_C6HC.
DR Pfam: PF01485; IIR; 2.
KW Hypothetical protein.
SQ SEQUENCE 497 AA; 57341 MW; 157B61283A58656F CRC64;

Query Match
81.6%; Score 31; DB 5; Length 497;

```

Best Local Similarity 77.8%; Pred. No. 2.9e+02;  
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 VLXDDLEA 9  
11111111  
Db 55 VLDDLEA 63

## RESULT 16

OQ3GY8 PRELIMINARY; PRT; 551 AA.  
AC OQ3GY8;  
DT 01-DEC-2001 (TREMBlrel. 19, Created)  
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
DE Non-ribosomal peptide synthetase.  
OS Streptomyces avermitilis.  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Streptomycetaceae; Streptomyces.  
OX NCBI\_TaxID=33903;

## SEQUENCE FROM N.A.

RA MEDLINE=21477403; PubMed=11572948;  
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,  
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,  
RA Kikuchi H., Shiba T., Sakai Y., Hattori M.;  
RT "Genome sequence of an industrial microorganism Streptomyces  
RT avermitilis. Deducing the ability of producing secondary  
RT metabolites.";  
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).  
DR EMBL: AB070956; BAB69403.1; -;  
DR InterPro: IPR001242; Condensatu.  
DR InterPro: IPR003880; Pantne\_atfch.  
DR Pfam: PF00668; Condensation; 1.  
DR Pfam: PF00550; PP-binding; 1.  
DR PROSITE: PSS0075; ACP\_DOMAIN; 1.  
DR Phosphopantetheine.  
SQ SEQUENCE 551 AA; 59104 MW; D0581F69F09BA95F CRC64;

Query Match 81.6%; Score 31; DB 2; Length 551;  
Best Local Similarity 77.8%; Pred. No. 3.3e+02;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 VLXDDLEA 9  
11111111  
Db 254 VLDDLEA 262

## RESULT 17

OQKD07 PRELIMINARY; PRT; 750 AA.

AC OQKD07;  
DT 01-OCT-2000 (TREMBlrel. 15, Created)  
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
DE Hypothetical protein BH1154.

## GN BH1154.

OS Bacillus halodurans.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;  
OC Bacillaceae; Bacillus.  
OX NCBI\_TaxID=86665;

## SEQUENCE FROM N.A.

RA STRAIN=C-125 / JCM 9153;  
RA MEDLINE=20512562; PubMed=11058132;  
RA Takami H., Nakasone K., Takai Y., Maeno G., Sasaki R., Masui N.,  
RA Fujii F., Hitama C., Nakamura Y., Ogasawara N., Kuhara S.,  
RA Horikoshi K.;  
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus  
RT halodurans and genomic sequence comparison with Bacillus subtilis.";  
RL Nucleic Acids Res. 28:4317-4331(2000).  
CC -1- SIMILARITY: TO PROKARYOTE SENSORY TRANSDUCTION PROTEINS.  
DR EMBL: AP001511; BAB04873.1; -;

DR HSP: P02933; 130Y.  
DR InterPro: IPR003594; ATPbind\_ATPase.  
DR InterPro: IPR003660; HAMP.  
DR InterPro: IPR003661; H1S\_KinA.  
DR InterPro: IPR004359; H1S\_Kin\_sig.  
DR Pfam: PF02518; HATPase\_C; 1.  
DR Pfam: PF00512; signal; 1.  
DR SMART: SM00304; HAMP; 1.  
DR SMART: SM00387; HATPase\_C; 1.  
DR SMART: SM00388; H1SKA; 1.  
DR Kinase; Phosphorylation; Sensory transduction; Transferase;  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 750 AA; 87283 MW; 262B7F2D1B9833B3 CRC64;

Query Match 81.6%; Score 31; DB 16; Length 750;  
Best Local Similarity 66.7%; Pred. No. 4.5e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 VLXDDLEA 9  
11111111  
Db 581 ILDDLEA 589

## RESULT 18

OQUB20 PRELIMINARY; PRT; 1042 AA.

AC OQUB20;  
DT 01-MAY-2000 (TREMBlrel. 13, Created)  
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
DE FS(1)YB.  
GN FS(1)YB OR FS OR EG:95B7.8 OR CG2706.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;

## SEQUENCE FROM N.A.

RA MEDLINE=99203434; PubMed=10101118;  
RA King F.J., Iln H.;  
RT "Somatic signaling mediated by fs(1)YB is essential for germline stem  
RT cell maintenance during Drosophila oogenesis.";  
RL Development 126:1833-1844(1999).  
DR EMBL: AF141673; AAB32687.1; -;  
DR FlyBase: FBgn0000928; fs(1)YB.  
SQ SEQUENCE 1042 AA; 117465 MW; D2B9F14F7642DBEF CRC64;

Query Match 81.6%; Score 31; DB 5; Length 1042;  
Best Local Similarity 77.8%; Pred. No. 6.4e+02;  
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 VLXDDLEA 9  
11111111  
Db 595 VLDDLEA 603

## RESULT 19

O76911 PRELIMINARY; PRT; 1042 AA.

AC O76911;  
DT 01-NOV-1998 (TREMBlrel. 08, Created)  
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)  
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
DE EG:95B7.8 protein.  
GN FS(1)YB OR EG:95B7.8 OR CG2706.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;

RP SEQUENCE FROM N.A.

RA Ferraz C., Vidal S., Brun C., Bucheton A., Demallie J.G.;  
 RT "Sequencing the distal X chromosome of *Drosophila melanogaster*."  
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
 RP [2]  
 RN SEQUENCE FROM N.A.  
 RA Benos P.;  
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL021728; CA16820.1;  
 DR FlyBase; FBgn0000928; fs(1)Yb.  
 SO SEQUENCE 1042 AA; 117450 MW; A2738A78ACB1DEAF CRC64;

Query Match  
 Best Local Similarity 81.6%; Score 31; DB 5; Length 1042;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 VLXDDLEA 9  
 Db 595 VLFPDLEA 603

RESULT 20  
 ID 09M4W2 PRELIMINARY; PRT: 1042 AA.  
 AC 09M4W2;  
 DT 01-MAY-2000 (TREMblrel. 13, Created)  
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)  
 DT 01-MAR-2001 (TREMblrel. 16, Last annotation update)  
 DE FS(1)Yb protein.  
 GN FS(1)Yb OR FS OR EG:95B7.8 OR CG2706.  
 OS *Drosophila melanogaster* (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BERKELEY;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Balow D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktoglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkov D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA DePamphilis M., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrera S., Fleischmann W.,  
 RA Folsler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodde A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibeagwa C.,  
 RA Jalili M., Kalish F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasok P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Mlshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclab J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reibert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skipski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svitskas R., Tector C., Turner R., Venter J., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Messerman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,

RA Gidys R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of *Drosophila melanogaster*."  
 RL Science 287:2185-2195(2000).  
 DR EMBL; AE003425; AAP45815.1;  
 DR FlyBase; FBgn0000928; fs(1)Yb.  
 SO SEQUENCE 1042 AA; 117451 MW; DF949D2BD80F1902 CRC64;

Query Match  
 Best Local Similarity 81.6%; Score 31; DB 5; Length 1042;  
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 VLXDDLEA 9  
 Db 595 VLFPDLEA 603

RESULT 21  
 ID 09XWD3 PRELIMINARY; PRT: 72 AA.  
 AC 09XWD3;  
 DT 01-NOV-1999 (TREMblrel. 12, Created)  
 DT 01-NOV-1999 (TREMblrel. 12, Last sequence update)  
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)  
 DE Y47H9C.12 protein.  
 GN Y47H9C.12.  
 OS *Caenorhabditis elegans*.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Harris B.R.;  
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99069613; PubMed=9851916;  
 RA none;  
 RT "Genome sequence of the nematode *C.elegans*: A platform for  
 RT investigating biology."  
 RL Science 282:2012-2018(1998).  
 DR EMBL; AL032657; CAA21742.1;  
 SO SEQUENCE 72 AA; 8379 MW; 8C0FDFD4E40E817 CRC64;

Query Match  
 Best Local Similarity 78.9%; Score 30; DB 5; Length 72;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLXDDLEA 8  
 Db 61 VLGDLEA 68

RESULT 22  
 ID 09H653 PRELIMINARY; PRT: 162 AA.  
 AC 09H653;  
 DT 01-MAR-2001 (TREMblrel. 16, Created)  
 DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)  
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)  
 DE CDNA: FLJ22601 f1s, clone HS104471.  
 OS *Homo sapiens* (Human).  
 OC Eukaryota; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=SMALL INTESTINE;  
 RA Watanabe K., Kumagai A., Itakura S., Yamazaki M., Tashiro H., Ota T.,  
 RA Suzuki Y., Oobayashi M., Nishi T., Shibahara T., Tanaka T.,  
 RA Nakamura Y., Isegaki T., Sugano S.,  
 RT "NEO human cDNA sequencing project."  
 RT Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AK026254; BAB15413.1;

DR InterPro: IPR002048; EF-hand.  
SQ SEQUENCE 162 AA; 19013 MW; E66AF08812A388A3 CRC64;

Query Match 78.9%; Score 30; DB 4; Length 162;  
Best Local Similarity 75.0%; Pred. No. 1.5e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 LXDLLLEA 9  
| | | | |  
DB 83 LSDLLLES 90

RESULT 23

ID 050241 PRELIMINARY; PRT; 217 AA.

AC 050241:  
DT 01-JUN-1998 (TREMBlrel. 06, Created)  
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)  
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
DE Transposase homolog.  
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).  
OC Plasmid pTIC58.  
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
OC Rhizobiaceae; Rhizobium.  
OX NCBI\_TaxID=176299;  
RN [1]  
RP SEQUENCE FROM N.A.

RC STRAIN=C58;  
RX MEDLINE=99009000; PubMed=9791116;  
RA Kallagher V.S., Wians S.C.;

RT "Wound-released chemical signals may elicit multiple responses from an  
RT Agrobacterium tumefaciens strain containing an octopine-type Ti  
RT plasmid."  
DR J. Bacteriol. 180:5660-5667(1998).  
DR EMBL; AF034769; AAC71784.1; -.

DR Plasmid.  
SQ SEQUENCE 217 AA; 24511 MW; 81AAB02DA7D2418D CRC64;

Query Match 78.9%; Score 30; DB 2; Length 217;  
Best Local Similarity 75.0%; Pred. No. 2e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 LXDLLLEA 9  
| | | | |  
DB 70 LTDVLEA 77

RESULT 24

ID 09ASD4 PRELIMINARY; PRT; 248 AA.

AC 09ASD4:  
DT 01-JUN-2001 (TREMBlrel. 17, Created)  
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)  
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
DE P0456F08.25 protein (B142C05.7 protein).  
GN P0456F08.25 OR B142C05.7.  
OS Oryza sativa (Rice), and  
OS Oryza sativa (japonica cultivar-group).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhartoideae; Oryzaceae; Oryza.  
OX NCBI\_TaxID=4530, 39947;

RP SEQUENCE FROM N.A.  
RC STRAIN=CV, NIPPONBARE;  
RA Sasaki T., Matsumoto T., Yamamoto K.;

RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC  
RT clone: P0456F08."  
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.  
RN [2]

RP SEQUENCE FROM N.A.  
RC STRAIN=CV, NIPPONBARE;  
RA Sasaki T., Matsumoto T., Yamamoto K.;

RT "Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 1, BAC  
RT clone: B142C05."  
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AP002901; BAB39425.1; -.  
DR EMBL; AP003410; BAB89107.1; -.  
DR HSSP; P36610; 1681.  
DR InterPro: IPR002048; EF-hand.  
DR InterPro: IPR001125; Recoverin.  
DR Pfam: PF00036; ehand; 3.  
DR PRINTS; P00450; RECOVERIN.  
DR ProDom; PD000012; EF-hand; 1.  
DR SMART; SM00054; EFn; 3.  
DR PROSITE; PS00018; EF-HAND; UNKNOWN\_1.

SQ SEQUENCE 248 AA; 27764 MW; 473B5A677612678 CRC64;

Query Match 78.9%; Score 30; DB 10; Length 248;  
Best Local Similarity 75.0%; Pred. No. 2.3e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 LXDLLLEA 9  
| | | | |  
DB 149 LSDVLEA 156

RESULT 25

ID 09N3G0 PRELIMINARY; PRT; 253 AA.

AC 09N3G0:  
DT 01-OCT-2000 (TREMBlrel. 15, Created)  
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
DE Hypothetical 27.0 kDa protein.  
GN Y53G8AR.6.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;  
OC Rhabditidae; Pelodierinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BRISTOL N2;  
RX MEDLINE=99069613; PubMed=9851916;

RA None;  
RT "Genome sequence of the nematode C. elegans: a platform for  
RT investigating biology. The C. elegans Sequencing Consortium."  
RL Science 282:2012-2018(1998).  
RN [2]

RP SEQUENCE FROM N.A.  
RC STRAIN=BRISTOL N2;  
RA Latreille P.;  
RT "The sequence of C. elegans cosmid Y53G8AR."  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
RN [3]

RP SEQUENCE FROM N.A.  
RC STRAIN=BRISTOL N2;  
RA Waterston R.;  
RT "Direct Submision."  
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AC024809; AAF59543.1; -.

DR InterPro: IPR003034; SAP.  
DR Pfam: PF02037; SAP; 1.  
DR SMART; SM00513; SAP; 1.  
KW Hypothetical protein.

SQ SEQUENCE 253 AA; 26960 MW; F573AD88CD18DA1 CRC64;

Query Match 78.9%; Score 30; DB 5; Length 253;  
Best Local Similarity 55.6%; Pred. No. 2.4e+02;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 LXDLLLEA 9  
| | | | |  
DB 66 LINDLLDA 74

```

RESULT 26
Q96BD2      PRELIMINARY;      PRT;      302 AA.
ID Q96BD2;
AC Q96BD2;
DT 01-MAY-2000 (TRENBLREL. 13, Created)
DT 01-MAY-2000 (TRENBLREL. 13, Last sequence update)
DT 01-DEC-2001 (TRENBLREL. 19, Last annotation update)
DE TIOF128 protein.
GN TIOF128.
OS Agrobacterium tumefaciens.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=358;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF301001;
RX MEDLINE=20184752; PubMed=10721727;
RA Suzuki K., Hattori Y., Uraji M., Ohta N., Iwata K., Murata K.,
RA Katoh A., Yoshida K.;
RT "Complete nucleotide sequence of a plant tumor-inducing Ti plasmid.";
RL Gene 242:331-336(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF301001;
RX MEDLINE=98193120; PubMed=9524202;
RA Suzuki K., Ohta N., Hattori Y., Uraji M., Katoh A., Yoshida K.;
RT "Novel structural difference between nopaline- and octopine- type trbJ
RT gene: construction of genetic and physical map and sequencing of
RT trbJ/trai and rep gene clusters of a new Ti plasmid pTi-SAKURA.";
RL Biochim. Biophys. Acta 1396:1-7(1998).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF301001;
RA Hattori Y., Suzuki K., Ohta N., Uraji M., Katoh A., Yoshida K.;
RT "Genome structure of pTi-SAKURA(I): Strategy for DNA sequencing of a
RT Japanese cherry-Ti plasmid.";
RL Nucleic Acids Symp. Ser. 37:159-160(1998).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF301001;
RA Ohta N., Suzuki K., Hattori Y., Uraji M., Katoh A., Yoshida K.;
RT "Genome structure of pTi-SAKURA (II): Characteristics of T-DNA.";
RL Nucleic Acids Symp. Ser. 39:185-186(1998).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF301001;
RA Uraji M., Suzuki K., Ohta N., Hattori Y., Katoh A., Yoshida K.;
RT "Genome structure of pTi-SAKURA (IV): Characteristics of tra region.";
RL Nucleic Acids Symp. Ser. 39:187-188(1998).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF301001;
RA Hattori Y., Suzuki K., Ohta N., Uraji M., Katoh A., Yoshida K.;
RT "Genome structure of pTi-SAKURA (V): Complete nucleotide sequence of
RT plasmid pTi-SAKURA's vir region in Agrobacterium tumefaciens.";
RL EMBL: AB016260; BA887753.1; -.
DR EMBL: AB016260; BA887753.1; -.
KW Plasmid.
SQ SEQUENCE 302 AA; 33571 MW; C8A67426139195F4 CRC64;

Query Match      78.9%; Score 30; DB 2; Length 302;
Best Local Similarity 75.0%; Pred. No. 2.8e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LXDLLLEA 9
   | | | | |
Db 59 LTDDVLEA 66

RESULT 27
Q805Z9      PRELIMINARY;      PRT;      302 AA.
ID Q805Z9

```

```

AC Q805Z9;
DT 01-JUN-2002 (TRENBLREL. 21, Created)
DT 01-JUN-2002 (TRENBLREL. 21, Last sequence update)
DT 01-JUN-2002 (TRENBLREL. 21, Last annotation update)
DE Transposase.
GN TNP OR AT06148 OR AGR_PTI-BX53.
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608550; PubMed=11743193;
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F., Jr., Woo L.,
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Boye D., St.,
RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA Kutayavin T., Levy R., Li M.-J., McClelland E., Palmieri A.,
RA Raymond C., Rouse G., Saenphimachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Nester E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
RT C58.";
RL Science 294:2317-2323(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608551; PubMed=11743194;
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RA Qurollo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
RA Houmel K., Gordon J., Vaudin M., Tartchouk O., Epp A., Liu F.,
RA Wollam C., Allinger M., Dougherty D., Scott C., Lappas C., Markelz B.,
RA Flanagan C., Crowell C., Gurson J., Lomo C., Sear C., Strub G.,
RA Cielo C., Slater S.;
RT "Genome sequence of the plant pathogen and biotechnology agent
RT Agrobacterium tumefaciens C58.";
RL Science 294:2323-2328(2001).
DR EMBL: AE009433; AAL46384.1; -.
DR EMBL: AE007940; AAK9112.1; -.
KW Plasmid; Complete proteome.
SQ SEQUENCE 302 AA; 33538 MW; BEAF932439195E0 CRC64;

Query Match      78.9%; Score 30; DB 16; Length 302;
Best Local Similarity 75.0%; Pred. No. 2.8e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LXDLLLEA 9
   | | | | |
Db 59 LTDDVLEA 66

RESULT 28
Q93J73      PRELIMINARY;      PRT;      327 AA.
ID Q93J73;
AC Q93J73;
DT 01-DEC-2001 (TRENBLREL. 19, Created)
DT 01-DEC-2001 (TRENBLREL. 19, Last sequence update)
DT 01-JUN-2002 (TRENBLREL. 21, Last annotation update)
DE Putative NAD-binding protein.
GN SC02824 OR SCBAC17F8.15.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycetales; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA O'Neil S., Harris D.;
RT Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.

```

RC STRAIN-A3(2); Parkhill J., Barrell B.G., Rajandream M.A.;  
 RA Bentley S.D., Parkhill J., Barrell B.G., Rajandream M.A.;  
 RL Submitted (Jul-2001) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A3(2);  
 RX MEDLINE-97000351; PubMed-8643436;  
 RA Redenbach M., Kleser H.M., Denapalte D., Eichner A., Cullum J.,  
 RA Kinashi H., Hopwood D.A.;  
 RT "A set of ordered cosmids and a detailed genetic and physical map for  
 the 8 Mb Streptomyces coelicolor A3(2) chromosome.";  
 RL Mol. Microbiol. 21:77-96(1996).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A3(2) / M145;  
 RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,  
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kleser H.,  
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,  
 RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,  
 RA Huang C.-H., Kleser T., Latke L., Murphy L., Oliver K., O'Neill S.,  
 RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,  
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,  
 RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,  
 RA Hopwood D.A.;  
 RT "Complete genome sequence of the model actinomycete Streptomyces  
 coelicolor A3(2)." ;  
 RL Nature 417:141-147(2002).  
 DR EMBL: AL596030; CAC44287.1; -;  
 DR InterPro: IPR002162; D\_2hyd\_dh.1.  
 DR Pfam: PF02826; 2-Hacid\_DH\_C.1.  
 SQ SEQUENCE 327 AA; 34874 MW; 46F9FE02F62F28F CRC64;

Query Match 78.9%; Score 30; DB 16; Length 327;  
 Best Local Similarity 66.7%; Pred. No. 3.1e+02;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 VLXDDLLEA 9  
 1:|||||  
 DB 240 VVEDLVEA 248

RESULT 29  
 ID 09M102 PRELIMINARY; PRT; 361 AA.  
 AC 09M102;  
 DT 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
 DE Serine/threonine protein kinase-like protein.  
 GN T17J13.180.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eustosids II; Brassicales; Brassicaceae; Arabidopsids.  
 OX NCBI\_Taxid=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Rieger M., Mueller-Auer S., Zipp M., Schaefer M., Mewes H.W.,  
 RA Lemcke K., Mayer K.F.X., Queller F., Salanoubat M.;  
 RL Submitted (Feb-2000) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA EU Arabidopsis sequencing project;  
 RL Submitted (Feb-2000) to the EMBL/GenBank/DBJ databases.  
 CC -1 STABILITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.  
 DR EMBL: AL138651; CAB71882.1; -;  
 DR InterPro: IPR000719; Euk\_kinase.  
 DR InterPro: IPR002290; Ser\_thr\_kinase.  
 DR InterPro: IPR004040; STY\_kinase.  
 DR InterPro: IPR001245; Tyr\_kinase.  
 DR Pfam: PF00069; kinase\_1.  
 DR PRINTS: PR00109; TYRKINASE.  
 DR ProDom: PD000001; Euk\_kinase; 1.

DR SMART: SM00221; STYKC; 1.  
 DR PROSITE: PS00107; PROTEIN\_KINASE\_ARP; 1.  
 DR PROSITE: PS50011; PROTEIN\_KINASE\_DOM; 1.  
 DR PROSITE: PS00108; PROTEIN\_KINASE\_ST; 1.  
 KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.  
 SQ SEQUENCE 361 AA; 38994 MW; 73EFE078A1F498FD CRC64;

Query Match 78.9%; Score 30; DB 10; Length 361;  
 Best Local Similarity 55.6%; Pred. No. 3.4e+02;  
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 VLXDDLLEA 9  
 1:|||||  
 DB 57 ILADELIEA 65

RESULT 30  
 ID 09HLH2 PRELIMINARY; PRT; 401 AA.  
 AC 09HLH2;  
 DT 01-MAR-2001 (TREMBlrel. 16, Created)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Hypothetical protein Ta0256.  
 GN TA0256.  
 OS Thermoplasma acidophilum.  
 OC Archaea; Euryarchaeota; Thermoplasmata; Thermoplasmatales;  
 OC Thermoplasmataceae; Thermoplasma.  
 OX NCBI\_Taxid=2303;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=DSM 1728;  
 RX MEDLINE-20479972; PubMed-11029001;  
 RA Ruepp A., Graml W., Santos-Martinez M.-L., Koretke K.K., Volker C.,  
 RA Mewes H.-W., Fishman D., Stocker S., Lupas A.N., Baumgaertel W.;  
 RT "The genome sequence of the thermophilic scavenger Thermoplasma  
 acidophilum." ;  
 RL Nature 407:508-513(2000).  
 DR EMBL: AL445063; CAC11401.1; -;  
 DR InterPro: IPR002792; TRAM.  
 DR InterPro: IPR001861; UPF0004.  
 DR Pfam: PF01938; TRAM; 1.  
 DR Pfam: PF00919; UPF0004; 1.  
 DR TIGRFAMs: TIGR00089; UPF0004; 1.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 401 AA; 45365 MW; 30AB6D61AE8D7ABA CRC64;

Query Match 78.9%; Score 30; DB 17; Length 401;  
 Best Local Similarity 75.0%; Pred. No. 3.8e+02;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 LXXDDLLEA 9  
 1:|||||  
 DB 209 ILDDLLEA 216

RESULT 31  
 ID 09ESF8 PRELIMINARY; PRT; 589 AA.  
 AC 09ESF8;  
 DT 01-MAR-2001 (TREMBlrel. 16, Created)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Inhibitor of apoptosis protein 2.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_Taxid=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Holcik M., Lefebvre C.A., Hicks K., Korneluk R.G.;  
 RT "Cloning and Characterization of the Rat Homologs of the Inhibitor of  
 Apoptosis Protein 1, 2, and 3 Genes." ;

Submitted (SEP-1999) to the EMBL/Genbank/DBJ databases.  
CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.

DR EMBL: AF183431; AAG22971.1; -  
DR HSSP: Q13490; 1QNH.  
DR InterPro: IPR001370; BIR.  
DR InterPro: IPR001315; CARD.  
DR InterPro: IPR001841; Znf\_fing.  
DR Pfam: PF00653; BIR; 3.  
DR Pfam: PF00619; CARD; 1.  
DR Pfam: PF00097; zf-C3HC4; 1.  
DR SMART: SM00238; BIR; 3.  
DR SMART: SM00114; CARD; 1.  
DR SMART: SM00184; RING; 1.  
DR PROSITE: PS01282; BIR\_REPEAT\_1; 2.  
DR PROSITE: PS0143; BIR\_REPEAT\_2; 3.  
DR PROSITE: PS0209; CARD; 1.  
KW Zinc-finger.  
SQ SEQUENCE 589 AA; 66777 MW; B4F7089BD7CD285B CRC64;

Query Match 78.9%; Score 30; DB 11; Length 589;  
Best Local Similarity 75.0%; Pred. No. 5.7e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 LXDDLLEA 9  
: |||||  
DB 449 IJDDLLEA 456

RESULT 32

O9QZC6 PRELIMINARY; PRT; 589 AA.

DT 01-MAY-2000 (TREMblrel. 13, Created)  
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)  
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)  
DE Inhibitor of apoptosis protein 2.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=KIDNEY;  
RA Dong Z., Denton M., Gu S.M., Saikumar P., Venkatchalam M.A.;  
RT Cloning of cDNA for rat inhibitor of apoptosis protein 2.;  
RT Submitted (SEP-1999) to the EMBL/Genbank/DBJ databases.  
CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.  
DR EMBL: AF190020; AAF04585.1; -  
DR HSSP: Q13490; 1QNH.  
DR InterPro: IPR001370; BIR.  
DR InterPro: IPR001315; CARD.  
DR InterPro: IPR001841; Znf\_fing.  
DR Pfam: PF00653; BIR; 3.  
DR Pfam: PF00097; CARD; 1.  
DR SMART: SM00238; BIR; 3.  
DR SMART: SM00114; CARD; 1.  
DR SMART: SM00184; RING; 1.  
DR PROSITE: PS01282; BIR\_REPEAT\_1; 3.  
DR PROSITE: PS0143; BIR\_REPEAT\_2; 3.  
DR PROSITE: PS0209; CARD; 1.  
KW Zinc-finger.  
SQ SEQUENCE 589 AA; 66777 MW; E6812FEF3EA34142 CRC64;

Query Match 78.9%; Score 30; DB 11; Length 589;  
Best Local Similarity 75.0%; Pred. No. 5.7e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 LXDDLLEA 9  
: |||||  
DB 449 IJDDLLEA 456

RESULT 33

O9VIX8 PRELIMINARY; PRT; 623 AA.

ID O9VIX8  
AC O9VIX8;  
DT 01-MAY-2000 (TREMblrel. 13, Created)  
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)  
DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)  
DE Type I restriction modification enzyme, subunit M.  
GN PAB2149.  
OS Pyrococcus abyssi.  
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;  
OC Pyrococcus.  
OX NCBI\_TaxID=29292;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ORSAY;  
RA Hellig R.;  
RT "Pyrococcus abyssi genome sequence: insights into archaeal chromosome structure and evolution."  
RT Submitted (JUL-1999) to the EMBL/Genbank/DBJ databases.  
DR EMBL: AJ248283; CAB49220.1; -  
DR InterPro: IPR003665; Methylase\_M.  
DR InterPro: IPR002296; N12N6\_mifrase.  
DR InterPro: IPR003356; N6\_DNA\_mtease.  
DR InterPro: IPR002052; N6\_mtease.  
DR Pfam: PF02506; Methylase\_M; 1.  
DR Pfam: PF02384; N6\_mtease; 1.  
DR PRINTS: PR00507; N12N6MTEFRASE.  
DR PROSITE: PS00092; N6\_MTEASE; UNKNOWN\_1.  
KW Complete proteome.  
SQ SEQUENCE 623 AA; 72679 MW; B87C455390DB8A03 CRC64;

Query Match 78.9%; Score 30; DB 17; Length 623;  
Best Local Similarity 55.6%; Pred. No. 6.1e+02;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 IJDDLLEA 9  
: |||||  
DB 472 IVEDDLLEA 480

RESULT 34

O9PD83 PRELIMINARY; PRT; 683 AA.

ID O9PD83  
AC O9PD83;  
DT 01-OCT-2000 (TREMblrel. 15, Created)  
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)  
DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)  
DE TonB-dependent receptor for iron transport.  
GN Xrl496.  
OS Xylella fastidiosa.  
OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;  
OX Xylella.  
OX NCBI\_TaxID=2371;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=9A5C;  
RA MEDLINE=20365717; PubMed=10910347;  
RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,  
RA Alvaranga R., Alves L.M.C., Araya J.E., Bata G.S., Baptista C.S.,  
RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Brites M.R.S.,  
RA Bueno M.R.P., Colombo A.A., Camargo L.E.A., Carriro D.M., Carrer H.,  
RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,  
RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorry H.,  
RA Facinani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,  
RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,  
RA Garner M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,  
RA Ho P.L., Hobeisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,  
RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,  
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,  
RA Machado W.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,  
RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,  
RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,



RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,  
 RA Niani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,  
 RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,  
 RA Paixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pasquero J.B.,  
 RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,  
 RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,  
 RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,  
 RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,  
 RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsubako M.H.,  
 RA Vallada H., Van Sluys M.A., Verjowski-Almeida S., Vettore A.L.,  
 RA Zago M.A., Zaitz M., Meidanis J., Setubal J.C.,  
 RT "The genome sequence of the plant pathogen *Xylella fastidiosa*,"  
 RL Nature 406:151-159(2000).  
 DR EMBL: AE003979; AAF84305.1; -;  
 DR InterPro: IPR000531; TonB\_boxC.  
 DR Pfam: PF00593; TonB\_boxC.1.  
 KW Complete proteome.  
 SQ SEQUENCE 683 AA; 76437 MW; 72458D6F39AFFB4D CRC64;

Query Match 78.9%; Score 30; DB 16; Length 683;  
 Best Local Similarity 75.0%; Pred. No. 6.7e+02;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLXDDLE 8  
 |||||  
 Db 65 VLXDDLE 72

## RESULT 35

094GPF8 PRELIMINARY; PRT; 799 AA.

AC 01-DEC-2001 (TREMBlrel. 19, Created)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE Putative polyprotein, 5'-partial (Fragment).  
 GN OJ111.B1.1.  
 OS Oryza sativa (Rice).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzaceae; Oryza.  
 OX NCBI\_TaxID=4530;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV\_NIPPONBARE;  
 RA Buell C.R., Yuan Q., Ouyang S., Moffat K.S., Hill J.N., Gansberger K.,  
 RA Brenner M., Burgess S., Hance M., Shvartsbeyn M., Taitlin T.,  
 RA Riggs F., Hsiao J., Zismann V., Blunt S., Pal G., Vanaken S.E.,  
 RA Ullrich T.R., Feldblum T.V., Kalb E., Quackenbush J.,  
 RA Salzberg S.L., White O., Fraser C.M.,  
 RT "Oryza sativa chromosome 3 BAC OJ111.B1 genomic sequence,"  
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AC091247; AAK82434.1; -;  
 DR InterPro: IPR001969; Aspartate-site.  
 DR InterPro: IPR001584; Rye.  
 DR InterPro: IPR000477; RYase.  
 DR Pfam: PF00665; rye.1.  
 DR Pfam: PF00078; rye.1.  
 DR PROSITE: PS00141; ASP\_PROMPASE; UNKNOWN.1.  
 KW Polyprotein; RNA-directed DNA polymerase.  
 FT NON-TER 1  
 SQ SEQUENCE 799 AA; 89365 MW; 223230AF54F2A818 CRC64;

Query Match 78.9%; Score 30; DB 10; Length 799;  
 Best Local Similarity 66.7%; Pred. No. 7.9e+02;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLXDDLE 9  
 |||||  
 Db 90 VLXDDLE 98

## RESULT 36

089443 PRELIMINARY; PRT; 962 AA.

AC 089443;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
 DE HELICASE.  
 GN B962L.  
 OS African swine fever virus (ASFV).  
 OC Viruses; dsDNA viruses, no RNA stage; Asfarviridae; Asfivirus.  
 OX NCBI\_TaxID=10497;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BA71V;  
 RX MEDLINE=96036500; PubMed=7483270;  
 RA Yanez R.J., Rodriguez J.M., Nogal M.L., Yuste L., Enriquez C.,  
 RA Rodriguez J.F., Vinuela E.,  
 RT "Immune protection conferred by the baculovirus-related glycoprotein  
 RT of Thogoto virus (Orthomyxoviridae)."  
 RL Virology 208:249-278(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BA71V;  
 RX MEDLINE=94233765; PubMed=8178480;  
 RA La Vega I., Gonzalez A., Blasco R., Calvo V., Vinuela E.,  
 RA "Nucleotide sequence and variability of the inverted terminal  
 RT repetitions of African swine fever virus DNA."  
 RL Virology 201:152-156(1994).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BA71V;  
 RX MEDLINE=90219205; PubMed=2325203;  
 RA Gonzalez A., Calvo V., Almazan F., Almendral J.M., Ramirez J.C.,  
 RA La Vega I., Blasco R., Vinuela E.,  
 RT "Multigene families in African swine fever virus: family 360."  
 RL J. Virol. 64:2073-2081(1990).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BA71V;  
 RX MEDLINE=90219204; PubMed=2325202;  
 RA Almendral J.M., Almazan F., Blasco R., Vinuela E.,  
 RT "Multigene families in African swine fever virus: family 110."  
 RL J. Virol. 64:2064-2072(1990).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BA71V;  
 RX MEDLINE=91134988; PubMed=1994575;  
 RA Canacho A., Vinuela E.,  
 RT "Protein p22 of African swine fever virus: an early structural protein  
 RT that is incorporated into the membrane of infected cells."  
 RL Virology 181:251-257(1991).  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BA71V;  
 RA Almazan F., Murguia J.R., Rodriguez J.M., La Vega I., Vinuela E.,  
 RT Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.  
 RN [7]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BA71V;  
 RX MEDLINE=94187118; PubMed=8139051;  
 RA Rodriguez J.M., Yanez R.J., Pan R., Rodriguez J.F., Salas M.L.,  
 RA Vinuela E.,  
 RT "Multigene families in African swine fever virus: family 505."  
 RL J. Virol. 68:2746-2751(1994).  
 RN [8]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BA71V;  
 RX MEDLINE=93346971; PubMed=8393914;  
 RA Yanez R.J., Rodriguez J.M., Rodriguez J.F., Salas M.L., Vinuela E.,  
 RT "African swine fever virus thymidylate kinase gene: sequence and  
 RT transcriptional mapping."  
 RL J. Gen. Virol. 74:1633-1638(1993).  
 RN [9]

RP SEQUENCE FROM N.A.  
 RC STRAIN-BA71V;  
 RX MEDLINE=94065656; PubMed=8245848;  
 RA Alcamí A., Angulo A., Vinuela E.;  
 RT "Mapping and sequence of the gene encoding the African swine fever  
 RT virion protein of MRF 11500.";  
 RL J. Gen. Virol. 74:2317-2324(1993).  
 RN [10]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BA71V;  
 RX MEDLINE=93277388; PubMed=8503790;  
 RA Munoz M., Freije J.M., Salas M.L., Vinuela E., Lopez-Otin C.;  
 RT "Structure and expression in E. coli of the gene coding for protein  
 RT p10 of African swine fever virus.";  
 RL Arch. Virol. 130:93-107(1993).  
 RN [11]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BA71V;  
 RX MEDLINE=90357780; PubMed=2389555;  
 RA Blasco R., Lopez-Otin C., Munoz M., Bockamp E.O., Simon-Mateo C.,  
 RA Vinuela E.;  
 RT "Sequence and evolutionary relationships of African swine fever virus  
 RT thymidine kinase.";  
 RL Virology 178:301-304(1990).  
 RN [12]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BA71V;  
 RX MEDLINE=93281390; PubMed=8506138;  
 RA Yanez R.J., Boursnell M., Nogal M.L., Yuste L., Vinuela E.;  
 RT "African swine fever virus encodes two genes which share significant  
 RT homology with the two largest subunits of DNA-dependent RNA  
 RT polymerases.";  
 RL Nucleic Acids Res. 21:2423-2427(1993).  
 RN [13]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BA71V;  
 RX MEDLINE=93353606; PubMed=8102411;  
 RA Rodriguez J.M., Yanez R.J., Almazan F., Vinuela E., Rodriguez J.F.;  
 RT "African swine fever virus encodes a CD2 homology responsible for the  
 RT adhesion of erythrocytes to infected cells.";  
 RL J. Virol. 67:5312-5320(1993).  
 RN [14]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BA71V;  
 RX MEDLINE=94085774; PubMed=8262374;  
 RA Yanez R.J., Rodriguez J.M., Boursnell M., Rodriguez J.F., Vinuela E.;  
 RT "Two putative African swine fever virus helicases similar to yeast  
 RT 'DEAF' pre-mRNA processing proteins and vaccinia virus ATPases D1L  
 RT and D6R.";  
 RL Gene 134:161-174(1993).  
 RN [15]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BA71V;  
 RX MEDLINE=90223993; PubMed=2327074;  
 RA Lopez-Otin C., Freije J.M., Parra F., Mendez E., Vinuela E.;  
 RT "Mapping and sequence of the gene coding for protein p72, the major  
 RT capsid protein of African swine fever virus.";  
 RL Virology 175:477-484(1990).  
 RN [16]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BA71V;  
 RX MEDLINE=94123986; PubMed=8293992;  
 RA Rodriguez J.M., Yanez R.J., Rodriguez J.F., Vinuela E., Salas M.L.;  
 RT "The DNA polymerase-encoding gene of African swine fever virus:  
 RT sequence and transcriptional analysis.";  
 RL Gene 136:103-110(1993).  
 RN [17]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BA71V;  
 RX MEDLINE=93327788; PubMed=8335009;  
 RA Simon-Mateo C., Andres G., Vinuela E.;  
 RT "Polyprotein processing in African swine fever virus: a novel gene  
 RT expression strategy for a DNA virus.";

RL EMBO J. 12:2977-2987(1993).  
 RN [18]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BA71V;  
 RX MEDLINE=93233210; PubMed=8474154;  
 RA Prados F.J., Vinuela E., Alcamí A.;  
 RT "Sequence and characterization of the major early phosphoprotein p32  
 RT of African swine fever virus.";  
 RL J. Virol. 67:2475-2485(1993).  
 RN [19]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BA71V;  
 RX MEDLINE=92260660; PubMed=1583732;  
 RA Alcamí A., Angulo A., Lopez-Otin C., Munoz M., Freije J.M.,  
 RA Carrascosa A.L., Vinuela E.;  
 RT "Amino acid sequence and structural properties of protein p12, an  
 RT African swine fever virus attachment protein.";  
 RL J. Virol. 66:3860-3868(1992).  
 RN [20]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BA71V;  
 RX MEDLINE=93174976; PubMed=8438592;  
 RA Yanez R.J., Vinuela E.;  
 RT "African swine fever virus encodes a DNA ligase.";  
 RL Virology 193:531-536(1993).  
 RN [21]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BA71V;  
 RX MEDLINE=93174941; PubMed=8382399;  
 RA Pena L., Yanez R.J., Revilla Y., Vinuela E., Salas M.L.;  
 RT "African swine fever virus guanylyltransferase.";  
 RL Virology 193:319-328(1993).  
 RN [22]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BA71V;  
 RX MEDLINE=95159428; PubMed=7856088;  
 RA Simon-Mateo C., Freije J.M., Andres G., Lopez-Otin C., Vinuela E.;  
 RT "Mapping and sequence of the gene encoding protein p17, a major  
 RT African swine fever virus structural protein.";  
 RL Virology 206:1140-1144(1995).  
 RN [23]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BA71V;  
 RX MEDLINE=92263807; PubMed=1316688;  
 RA Garcia-Beato R., Freije J.M., Lopez-Otin C., Blasco R., Vinuela E.,  
 RA Salas M.L.;  
 RT "A gene homologous to topoisomerase II in African swine fever virus.";  
 RL Virology 188:938-947(1992).  
 RN [24]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BA71V;  
 RX MEDLINE=94091056; PubMed=8266720;  
 RA Freije J.M., Iain S., Vinuela E., Lopez-Otin C.;  
 RT "Nucleotide sequence of a nucleoside triphosphate phosphodiesterase  
 RT gene from African swine fever virus.";  
 RL  
 Qy 1 VLXDDLE 8  
 Db 584 LLADDDLE 591  
 Query Match 78.9%; Score 30; DB 12; Length 962;  
 Best local similarity 75.0%; Pred. No. 9,6e+02;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 RESULT 37  
 08V9U2  
 ID 08V9U2 PRELIMINARY; PRT; 963 AA.  
 AC 08V9U2;  
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE RNA helicase.

```

GN L09CL.
OS African swine fever virus (ASFV).
OC Viruses; dsDNA viruses, no RNA stage; Asfarviridae; Asfivirus.
OX NCBI_TaxID=10497;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MALAWI LIL-20/1;
RA Roberts P.C., Lu Z., Rock D.L.;
RT "Nucleotide sequence and analysis of 16.25 kilobase pairs of the
RT African swine fever virus genome that span the central variable
RT region."
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: L00966; AAL31320.1; -.
DR InterPro: IPR001410; DEAD.
DR InterPro: IPR002464; DEAD_box.
DR InterPro: IPR001650; Helicase_C.
DR Pfam: PF00271; helicase_C.1.
DR SMART: SM00487; DEXDC.1.
DR SMART: SM00490; HELICG.1.
DR PROSITE: PS00690; DEAH_ATP_HELICASE; UNKNOWN.1.
KW Helicase.
SQ SEQUENCE 963 AA; 109831 MW; C1137DAB22E2810F CRC64;

Query Match 78.9%; Score 30; DB 12; Length 963;
Best Local Similarity 75.0%; Pred. No. 9.6e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLXDDLE 8
DB 584 LIADDLE 591

RESULT 38
OBSAX4 PRELIMINARY; PRT; 1087 AA.
ID OBSAX4:
AC 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Putative polyprotein.
OS OSJNB0010E04.6.
GN Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Buell C.R., Yuan Q., Ouyang S., Liu J., Moffat K.S., Hill J.N.,
RA Gansberger K., Brenner M., Burgess S., Hance M., Shvartsbeyn M.,
RA Taitlin T., Riggs F., Hsiao J., Zisman Y., Blunt S., Pal G.,
RA Vanaken S.E., Ullerbach T.R., Feldlyum T.V., Kald E., Quackenbush J.,
RA Salzberg S.L., White O., Fraser C.M.;
RT "Oryza sativa chromosome 3 BAC OSJNB0010E04 genomic sequence.";
RT Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC096687; AAL79762.1; -.
KW Polyprotein.
SQ SEQUENCE 1087 AA; 121836 MW; A9B94AF2CC66139E CRC64;

Query Match 78.9%; Score 30; DB 10; Length 1087;
Best Local Similarity 66.7%; Pred. No. 1.1e+03;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLXDDLEA 9
DB 378 VLTDELDA 386

RESULT 39
OBSG71 PRELIMINARY; PRT; 1201 AA.
ID OBSG71:
AC 08RG71:

```

```

DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Hypothetical protein FN0446.
GN FN0446.
OS Fusobacterium nucleatum (subsp. nucleatum).
OC Bacteria; Fusobacteria; Fusobacterium.
OX NCBI_TaxID=76856;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 25586;
RX MEDLINE=21886394; PubMed=11889109;
RA Kapral V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,
RA Bhattacharya A., Barman A., Gardner W., Grechkin G., Zhu L.,
RA Vasileva O., Chu L., Kogan Y., Chagar O., Goldsman E., Bernal A.,
RA Larsen N., D'Souza M., Malunas T., Pusch G., Haselkorn R.,
RA Fongstein M., Kyrides N., Overbeek R.;
RT "Genome sequence and analysis of the oral bacterium Fusobacterium
RT nucleatum strain ATCC 25586."
RL J. Bacteriol. 184:2005-2018(2002).
DR EMBL: AF010556; AAL94642.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 1201 AA; 140781 MW; 3744CFB7261405E CRC64;

Query Match 78.9%; Score 30; DB 16; Length 1201;
Best Local Similarity 75.0%; Pred. No. 1.2e+03;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLXDDLE 8
DB 678 VLYDDLE 685

RESULT 40
OBSK19 PRELIMINARY; PRT; 1410 AA.
ID OBSK19:
AC 09K129:
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE EPOR.
GN EPOR.
OC Polyangium cellulosum.
OC Bacteria; Proteobacteria; delta subdivision; Myxobacteria;
OC Myxococcales; Sorangium; Polyangiaceae; Polyangium.
OX NCBI_TaxID=56;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SMp44;
RX MEDLINE=20293058; PubMed=10831849;
RA Julien B., Shah S., Ziemann R., Goldman R., Katz L., Khosla C.;
RT "Isolation and characterization of the epothilone biosynthetic gene
RT cluster from Sorangium cellulosum."
RL Gene 249:153-160(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=SMp44;
RX MEDLINE=20115953; PubMed=10649995;
RA Tang L., Shah S., Chung L., Carney J., Katz L., Khosla C., Julien B.;
RT "Cloning and heterologous expression of the epothilone gene cluster.";
RL Science 287:640-642(2000).
DR EMBL: AF217189; AAF62881.1; -.
DR HSSP: P14687; IANU.
DR InterPro: IPR000873; AMP-bind.
DR InterPro: IPR001242; Condensatn.
DR InterPro: IPR003880; Ppantine-attach.
DR Pfam: PF00501; AMP-binding.1.
DR Pfam: PF00668; Condensation.1.
DR Pfam: PF00550; pp-binding.1.
DR PRINTS: PR00154; AMPBINDING.
DR PROSITE: PS50075; ACP_DOMAIN.1.
DR PROSITE: PS00455; AMP_BINDING.1.
KW Phosphopantetheine.

```

SQ SEQUENCE 1410 AA; 158078 MW; C5C3780DC293B3AA CRC64;  
 Query Match  
 Best Local Similarity 78.9%; Score 30; DB 2; Length 1410;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 OY 2 LKXDDLLEA 9  
 DB 474 LKXDDLLEA 481  
 RESULT 41  
 ID Q9P8H3 PRELIMINARY; PRT; 1534 AA.  
 AC Q9P8H3;  
 DT 01-OCT-2000 (TREMUREL. 15, Created)  
 DT 01-OCT-2000 (TREMUREL. 15, Last sequence update)  
 DE 01-JUN-2002 (TREMUREL. 21, Last annotation update)  
 DE RecQ helicase MOSN.  
 GN MOSN.  
 OS Emericella nidulans (Aspergillus nidulans).  
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;  
 OC Eutotiales; Trichocomaceae; Emericella.  
 OX NCBI\_TaxID=5072;  
 RN 11  
 RP SEQUENCE FROM N.A.  
 RA Hotmann A.F., Harris S.D.;  
 RT "Suppression of ATM kinase defects by mutation of a RecQ helicase in  
 Aspergillus nidulans";  
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF259396; AAF72650.1; -;  
 DR Interpro: IPR001410; DEAD.  
 DR Interpro: IPR002464; DEAD\_box.  
 DR Interpro: IPR001650; Helicase\_C.  
 DR Interpro: IPR004589; RecQ.  
 DR Pfam: PF00270; DEAD; 1.  
 DR Pfam: PF00271; helicase\_C; 1.  
 DR SMART: SM00487; DEXDC; 1.  
 DR SMART: SM00490; HELIC; 1.  
 DR TIGRfam: TIGR00614; RecQ; 1.  
 DR PROSITE: PS00690; DEAD\_ATP\_HELICASE; UNKNOWN\_1.  
 KW ATP-binding; Helicase.  
 SQ SEQUENCE 1534 AA; 172538 MW; 705F53A7034A684 CRC64;  
 Query Match  
 Best Local Similarity 78.9%; Score 30; DB 3; Length 1534;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 OY 1 VLXDDLLEA 9  
 DB 625 VLXDDLLEA 633  
 RESULT 42  
 ID Q8YWB9 PRELIMINARY; PRT; 1999 AA.  
 AC Q8YWB9;  
 DT 01-MAR-2002 (TREMUREL. 20, Created)  
 DT 01-MAR-2002 (TREMUREL. 20, Last sequence update)  
 DE 01-MAR-2002 (TREMUREL. 20, Last annotation update)  
 DE Hypothetical protein ALL1696.  
 GN ALL1696.  
 OS Anabaena sp. (strain PCC 7120).  
 OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.  
 OX NCBI\_TaxID=103690;  
 RN 11  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=21595285; PubMed=11759840;  
 RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,  
 RA Watanabe A., Iriyuchi M., Ishikawa A., Kawashima K., Kimura T.,  
 RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,  
 RA Nakazaki N., Shimo S., Sugimoto M., Takazawa M., Yamada M.,  
 RA Yasuda M., Tabata S.;

RT "Complete genomic sequence of the filamentous nitrogen-fixing  
 cyanobacterium Anabaena sp. strain PCC 7120";  
 RL DNA Res. 8:205-213(2001).  
 DR EMBL: AF003586; BAB78062.1; -;  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 1999 AA; 207906 MW; DC0751D1BC4F1F53 CRC64;  
 Query Match  
 Best Local Similarity 78.9%; Score 30; DB 16; Length 1999;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 OY 2 LKXDDLLEA 9  
 DB 3 LKXDDLLEA 10  
 RESULT 43  
 ID Q87314 PRELIMINARY; PRT; 4976 AA.  
 AC Q87314;  
 DT 01-NOV-1998 (TREMUREL. 08, Created)  
 DT 01-NOV-1998 (TREMUREL. 08, Last sequence update)  
 DE 01-MAR-2002 (TREMUREL. 20, Last annotation update)  
 DE FxBC.  
 GN FxBC.  
 OS Mycobacterium smegmatis.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.  
 OX NCBI\_TaxID=1772;  
 RN 11  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MC2155;  
 RX MEDLINE=95191405; PubMed=7885224;  
 RA Fiss E.H., Yu S., Jacobs W.R. Jr.;  
 RT "Identification of genes involved in the sequestration of iron in  
 Mycobacteria: the ferric exochelin biosynthetic and uptake pathways";  
 RL Mol. Microbiol. 14:557-569(1994).  
 RN 12  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MC2155;  
 RX MEDLINE=98389687; PubMed=9721311;  
 RA Yu S., Fiss E., Jacobs W.R. Jr.;  
 RT "Analysis of the exochelin locus in mycobacterium smegmatis:  
 biosynthesis genes have homology with genes of the peptide synthetase  
 family";  
 RL J. Bacteriol. 180:4676-4685(1998).  
 DR EMBL: AF027770; AAC82550.1; -;  
 DR HSP; P14687; LAMU.  
 DR Interpro: IPR000873; AMP-bind.  
 DR Interpro: IPR000515; BPD\_transp.  
 DR Interpro: IPR001242; Condensatn.  
 DR Interpro: IPR000734; Lipase.  
 DR Interpro: IPR003880; Pantoate\_attch.  
 DR Interpro: IPR001031; Phosphatase.  
 DR Pfam: PF00501; AMP-binding; 4.  
 DR Pfam: PF00668; Condensation; 5.  
 DR Pfam: PF00550; pp-binding; 4.  
 DR PRINTS: PR00154; AMPBINDING.  
 DR PROSITE: PS00075; ACP\_DOMAIN; 4.  
 DR PROSITE: PS00402; BPD\_TRANS\_PAN\_MEMBER; UNKNOWN\_1.  
 DR PROSITE: PS00120; LIPASE\_SER; UNKNOWN\_1.  
 DR PROSITE: PS00012; PHOSPHOPANTETHEINE; UNKNOWN\_2.  
 KW Phosphopantetheine.  
 SQ SEQUENCE 4976 AA; 535912 MW; 3947A924441E03D CRC64;  
 Query Match  
 Best Local Similarity 78.9%; Score 30; DB 2; Length 4976;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 OY 1 VLXDDLLEA 9

Db 2308 ILDDLAE 2316

## RESULT 44

Q9JW27 PRELIMINARY; PRT; 78 AA.  
 AC Q9JW27;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE Hypothetical protein NMA0574.  
 GN NMA0574.  
 OS Neisseria meningitidis (serogroup A).  
 OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.  
 NCBI\_TaxID=5699;  
 RX MEDLINE=22491 / SEROGROUP A / SEROTYPE 4A;  
 RX MEDLINE=20222556; PubMed=10761919;  
 RA Pakhili J., Achtman M., James K.D., Bentley S.D., Churcher C., Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T., Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holtroyd S., Jagels K., Leather S., Moule S., Mungall K., Quail M.A., Ratandream M.A., Rutherford K.M., Simmonds M., Skelton J., Whitehead S., Spratt B.G., Barrall B.G.;  
 RT "Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.";  
 RL Nature 404:502-506(2000).  
 DR EMBL; AL162753; CAB83865.1; -;  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 78 AA; 8884 MW; A25602ACC1025941 CRC64;

Query Match 76.3%; Score 29; DB 16; Length 78;  
 Best Local Similarity 85.7%; Pred. No. 1.1e+02;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 LXDLE 8  
 ID 1  
 DB 61 LSDLE 67

RESULT 45  
 Q9JLB6 PRELIMINARY; PRT; 81 AA.  
 AC Q9JLB6;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
 DE Testis nuclear RNA-binding protein (Fragment).  
 GN IL2 OR TENR.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.  
 NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6;  
 RX MEDLINE=20243845; PubMed=10779485;  
 RA Lyons P.A., Amittage N., Argentina F., Denny P., Hill N.J., Lord C.J., Wilusz M.B., Peterson L.B., Wicker L.S., Todd J.A.;  
 RT "Congenic mapping of the type 1 diabetes locus, Id3, to a 780-kb region of mouse chromosome 3: identification of a candidate segment of ancestral DNA by haplotype mapping.";  
 RT Genome Res. 10:446-453(2000).  
 RL EMBL; AF195956; AAF32273.1; -;  
 DR MGD; MGI:103258; Tenr.  
 DR MGD; MGI:96548; 112.  
 DR InterPro; IPR002466; A\_deamin.  
 DR Pfam; PF02137; A\_deamin; 1.  
 DR PROSITE; PS50141; A\_DEAMIN\_EDIFASE; 1.  
 FT NON\_TER 1  
 DE SEQUENCE 81 AA; 9018 MW; OFE2FB5A8030CEB9 CRC64;

Query Match 76.3%; Score 29; DB 11; Length 81;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 DDLLE 9  
 ID 1  
 DB 32 DDLLE 37

## RESULT 46

Q8UEB3 PRELIMINARY; PRT; 101 AA.  
 AC Q8UEB3;  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE Hypothetical protein Atu1847.  
 GN Atu1847 OR AGR\_C\_3387.  
 OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).  
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group; Rhizobiaceae; Rhizobium.  
 NCBI\_TaxID=176299;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21608550; PubMed=11743193;  
 RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitaajima J.P., Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F., Jr., Woo L., Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr., Chapman P., Clendinning J., Deatherage G., Gillet W., Grant C., Kutyavin T., Levy R., Li M.-J., McClelland E., Palmieri A., Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D., Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M., Gordon-Kamm B., Lao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M., Chunley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V., Nester E.W.;  
 RT "The genome of the natural genetic engineer Agrobacterium tumefaciens C58.";  
 RL Science 294:2317-2323(2001).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21608551; PubMed=11743194;  
 RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M., Quirillo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L., Houmel K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F., Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B., Flanagan C., Crowell C., Gursion J., Lomo C., Sear C., Strub G., Cielo C., Slater S.;  
 RT "Genome sequence of the plant pathogen and biotechnology agent Agrobacterium tumefaciens C58.";  
 RL Science 294:2323-2328(2001).  
 DR EMBL; AE009139; AAL42843.1; -;  
 DR EMBL; AE008104; AAK87614.1; -;  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 101 AA; 11705 MW; E033AAZAAADDA25D CRC64;

Query Match 76.3%; Score 29; DB 16; Length 101;  
 Best Local Similarity 85.7%; Pred. No. 1.5e+02;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 LXDLE 8  
 ID 1  
 DB 69 LSDLE 75

RESULT 47  
 O53107 PRELIMINARY; PRT; 103 AA.  
 AC O53107;  
 DT 01-JUN-1998 (TrEMBLrel. 06, Created)  
 DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE Truncated Vaa surface lipoprotein adhesin (Fragment).  
 GN VAA.

OS Mycoplasma hominis.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;  
 OC Mycoplasmataceae; Mycoplasma.  
 OX NCBI\_TaxID=2098;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=183;  
 RX MEDLINE=98367128; PubMed=9701806;  
 RA Boesen T., Emmersen J., Jensen L.T., Ladefoged S.A., Thorsen P.,  
 RT Birkelund S.B., Christiansen G.,  
 RT "The Mycoplasma hominis vaa/p50 genes display a mosaic gene  
 structure."  
 RL Mol. Microbiol. 29:97-110(1998).  
 DR EMBL: AJ001667; CA04912.1; -  
 DR InterPro: IPR002520; Lipoprotein\_7.  
 DR Pfam: PF01540; Lipoprotein\_7; 1.  
 KW Lipoprotein.  
 FT NON\_TER 1  
 SQ SEQUENCE 103 AA: 1250 MW: D58564F5FA25AFA3 CRC64;

Query Match 76.3%; Score 29; DB 16; Length 103;  
 Best Local Similarity 75.0%; Pred. No. 1.5e+02;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 VLXDDLE 8  
 : | | | | |  
 DB 43 LVDLDLE 50

RESULT 48  
 O9X834 PRELIMINARY; PRT; 119 AA.  
 AC O9X834;  
 DT 01-NOV-1999 (TReMBLrel. 12, Created)  
 DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)  
 DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)  
 DE Hypothetical protein SC06067.  
 GN SC06067 OR SC9BL14C.  
 OS Streptomyces coelicolor.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.  
 OX NCBI\_TaxID=1902;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A3(2) / M145;  
 RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,  
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kleser H.,  
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,  
 RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,  
 RA Huang C.-H., Kleser T., Larke L., Murphy L., Oliver K., O'Neill S.,  
 RA Rabinowitz E., Rajandream M.A., Rutherford K., Rutter S.,  
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,  
 RA Warren T., Wietzorek A., Woodward J., Barrrell B.G., Parkhill J.,  
 RA Hopwood D.A.;  
 RT "Complete genome sequence of the model actinomycete Streptomyces  
 RL Nature 417:141-147(2002).  
 DR EMBL: AL049727; CAB41560.1; -  
 KW Hypothetical protein.  
 SQ SEQUENCE 119 AA: 12821 MW: AA5DBAD23E6BC065 CRC64;

Query Match 76.3%; Score 29; DB 16; Length 119;  
 Best Local Similarity 66.7%; Pred. No. 1.7e+02;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 VLXDDLEA 9  
 : | | | | |  
 DB 80 ILSDLEA 88

RESULT 49  
 O8R29 PRELIMINARY; PRT; 137 AA.  
 ID O8R29

AC O8R29;  
 DT 01-MAR-2002 (TReMBLrel. 20, Created)  
 DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)  
 DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)  
 DE Hypothetical protein ybdb.  
 GN ybdb OR STM059.  
 OS Salmonella typhimurium.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Salmonella.  
 OX NCBI\_TaxID=602;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=IT2 / SSGC1412 / ATCC 700720;  
 RX MEDLINE=21534948; PubMed=11677609;  
 RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,  
 RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,  
 RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,  
 RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,  
 RA Waterston R., Wilson R.K.;  
 RT "Complete genome sequence of Salmonella enterica serovar Typhimurium  
 RL Nature 413:852-856(2001).  
 DR EMBL: AE008723; AAL19550.1; -  
 DR InterPro: IPR003736; DUF157.  
 DR Pfam: PF02584; DUF157; 1.  
 DR TIGRFAMS: TIGR00369; unchar\_dom\_1; 1.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 137 AA: 14939 MW: 2C04E6D152BC5018 CRC64;

Query Match 76.3%; Score 29; DB 16; Length 137;  
 Best Local Similarity 75.0%; Pred. No. 2e+02;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 LXDLDLEA 9  
 : | | | | |  
 DB 31 LSGDDVLEA 38

RESULT 50  
 O8R29 PRELIMINARY; PRT; 137 AA.  
 AC O8R29;  
 DT 01-MAR-2002 (TReMBLrel. 20, Created)  
 DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)  
 DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)  
 DE Hypothetical protein STY0643.  
 GN STY0643.  
 OS Salmonella typhi.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Salmonella.  
 OX NCBI\_TaxID=601;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CT18;  
 RX MEDLINE=21534947; PubMed=11677608;  
 RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Main J.,  
 RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebatina M.,  
 RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,  
 RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,  
 RA Fellwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagers K.,  
 RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,  
 RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,  
 RA Whitehead S., Barrrell B.G.;  
 RT "Complete genome sequence of a multiple drug resistant Salmonella  
 RL Nature 413:848-852(2001).  
 DR EMBL: AL627267; CAD05075.1; -  
 DR InterPro: IPR003736; DUF157.  
 DR Pfam: PF02584; DUF157; 1.  
 DR TIGRFAMS: TIGR00369; unchar\_dom\_1; 1.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 137 AA: 14925 MW: 2C04E6D14B94F28B CRC64;

Query Match 76.3%; Score 29; DB 16; Length 137;  
 Best Local Similarity 75.0%; Pred. No. 2e+02;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 2 LKDDLEA 9  
 | | | | |  
 Db 31 LKDDVLEA 38

Search completed: December 27, 2002, 14:43:18  
 Job time : 71 secs





GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 27, 2002, 14:40:33 ; Search time 14 Seconds  
(Without alignments)  
18.915 Million cell updates/sec

Title: US-09-489-760A-1  
Perfect score: 38  
Sequence: 1 VLXDDLEA 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 1000 summaries

Database :

Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PTUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match length	ID	Description
1	32	84.2	4150 4 US-09-428-517-2	Sequence 2, Appli
2	30	78.9	1410 3 US-09-335-409-3	Sequence 3, Appli
3	30	78.9	1410 4 US-09-568-102-3	Sequence 3, Appli
4	30	78.9	1410 4 US-09-567-969-3	Sequence 3, Appli
5	30	78.9	1410 4 US-09-568-480-3	Sequence 3, Appli
6	30	78.9	1410 4 US-09-568-486-3	Sequence 3, Appli
7	30	78.9	1410 4 US-09-568-472-3	Sequence 3, Appli
8	30	78.9	1410 4 US-09-567-899-3	Sequence 3, Appli
9	28	73.7	251 5 PCT-US95-12357A-1	Sequence 1, Appli
10	28	73.7	550 1 US-08-674-168-29	Sequence 19, Appli
11	28	73.7	550 3 US-08-985-908-19	Sequence 4, Appli
12	28	73.7	550 3 US-08-852-730-4	Sequence 11, Appli
13	28	73.7	550 4 US-08-985-916-11	Sequence 4, Appli
14	28	73.7	3457 2 US-08-416-603-4	Sequence 4, Appli
15	28	73.7	3519 4 US-09-428-517-4	Sequence 3, Appli
16	28	73.7	3816 4 US-09-428-517-3	Sequence 6, Appli
17	27	71.1	111 4 US-08-899-330-6	Sequence 192, App
18	27	71.1	313 3 US-08-480-640A-192	Sequence 192, App
19	27	71.1	313 4 US-08-686-968C-192	Sequence 192, App
20	27	71.1	313 4 US-08-488-237A-192	Sequence 192, App
21	27	71.1	313 4 US-08-375-992A-192	Sequence 89, Appli
22	27	71.1	324 4 US-09-154-750A-192	Sequence 7, Appli
23	27	71.1	324 4 US-09-222-938A-31	Sequence 31, Appli
24	27	71.1	333 4 US-09-562-737-7	Sequence 114, App
25	27	71.1	389 3 US-08-480-640A-114	Sequence 114, App
26	27	71.1	389 3 US-08-295-802-114	Sequence 114, App
27	27	71.1	389 4 US-08-488-237A-114	Sequence 114, App

28	71.1	389 4 US-08-375-992A-114	Sequence 114, App
29	71.1	404 4 US-09-167-299-5	Sequence 5, Appli
30	71.1	748 4 US-09-409-180A-1	Sequence 1, Appli
31	71.1	709 5 PCT-US92-00731-13	Sequence 13, Appli
32	71.1	821 1 US-07-935-311A-4	Sequence 4, Appli
33	71.1	821 1 US-08-368-079-4	Sequence 4, Appli
34	71.1	821 5 PCT-US93-07996-4	Sequence 4, Appli
35	71.1	985 2 US-08-680-326-41	Sequence 41, Appli
36	71.1	1045 1 US-08-452-083-2	Sequence 2, Appli
37	68.4	32 3 US-08-851-843A-141	Sequence 141, App
38	68.4	32 4 US-08-974-549A-259	Sequence 259, App
39	68.4	32 4 US-08-854-050-141	Sequence 141, App
40	68.4	32 4 US-09-430-323-141	Sequence 141, App
41	68.4	35 4 US-08-974-549A-46	Sequence 46, Appli
42	68.4	49 3 US-08-851-843A-124	Sequence 124, App
43	68.4	49 4 US-08-974-549A-244	Sequence 244, App
44	68.4	49 4 US-08-854-050-124	Sequence 124, App
45	68.4	49 4 US-09-430-323-124	Sequence 124, App
46	68.4	69 3 US-08-851-843A-17	Sequence 17, Appli
47	68.4	69 4 US-08-854-050-17	Sequence 17, Appli
48	68.4	69 4 US-09-430-323-17	Sequence 17, Appli
49	68.4	241 4 US-09-153-599A-5	Sequence 5, Appli
50	68.4	291 4 US-09-134-001C-5669	Sequence 5669, Ap
51	68.4	314 2 US-08-408-095-23	Sequence 23, Appli
52	68.4	333 4 US-09-453-195A-2	Sequence 2, Appli
53	68.4	333 4 US-09-562-737-1	Sequence 1, Appli
54	68.4	333 4 US-09-562-737-5	Sequence 5, Appli
55	68.4	333 4 US-09-562-737-6	Sequence 6, Appli
56	68.4	333 4 US-09-562-737-8	Sequence 8, Appli
57	68.4	434 1 US-08-404-127-2	Sequence 2, Appli
58	68.4	434 2 US-09-041-011-2	Sequence 2, Appli
59	68.4	445 1 US-08-404-127-4	Sequence 4, Appli
60	68.4	445 2 US-09-041-011-4	Sequence 4, Appli
61	68.4	448 4 US-09-143-954-4	Sequence 4, Appli
62	68.4	448 4 US-09-323-195A-18	Sequence 18, Appli
63	68.4	457 4 US-09-143-954-2	Sequence 2, Appli
64	68.4	523 4 US-09-323-195A-17	Sequence 17, Appli
65	68.4	588 4 US-09-134-001C-3768	Sequence 3768, Ap
66	68.4	586 2 US-08-630-822A-70	Sequence 70, Appli
67	68.4	586 2 US-09-005-069-70	Sequence 70, Appli
68	68.4	586 4 US-09-171-156A-30	Sequence 30, Appli
69	68.4	659 4 US-09-562-737-11	Sequence 11, Appli
70	68.4	660 4 US-09-562-737-19	Sequence 19, Appli
71	68.4	660 3 US-08-819-177-1	Sequence 1, Appli
72	68.4	835 4 US-09-284-819-6	Sequence 6, Appli
73	68.4	868 4 US-09-398-395A-56	Sequence 56, Appli
74	68.4	884 3 US-08-851-843A-55	Sequence 55, Appli
75	68.4	884 4 US-08-974-549A-222	Sequence 222, Appli
76	68.4	884 4 US-08-854-050-55	Sequence 55, Appli
77	68.4	884 4 US-09-430-323-55	Sequence 55, Appli
78	68.4	1503 4 US-08-976-255-14	Sequence 14, Appli
79	68.4	2165 1 US-08-514-975B-2	Sequence 2, Appli
80	68.4	2165 5 PCT-US95-12507-2	Sequence 2, Appli
81	68.4	2594 4 US-08-718-388-7	Sequence 7, Appli
82	68.4	5405 4 US-08-718-388-9	Sequence 9, Appli
83	65.8	22 3 US-08-940-095-133	Sequence 133, App
84	65.8	22 3 US-08-940-093-133	Sequence 133, App
85	65.8	22 3 US-08-940-096-133	Sequence 133, App
86	65.8	22 4 US-09-465-719-133	Sequence 133, App
87	65.8	22 4 US-09-453-605-133	Sequence 133, App
88	65.8	22 4 US-09-453-838-133	Sequence 133, App
89	65.8	60 3 US-08-461-607-11	Sequence 11, Appli
90	65.8	60 4 US-09-363-600-11	Sequence 11, Appli
91	65.8	80 3 US-08-461-607-5	Sequence 5, Appli
92	65.8	80 4 US-09-363-600-5	Sequence 5, Appli
93	65.8	117 4 US-08-961-083-128	Sequence 128, App
94	65.8	118 4 US-09-413-814-17	Sequence 17, Appli
95	65.8	154 4 US-08-858-207A-433	Sequence 433, App
96	65.8	180 1 US-08-447-010-8	Sequence 8, Appli
97	65.8	181 4 US-08-961-083-62	Sequence 62, Appli
98	65.8	221 4 US-09-247-373B-54	Sequence 54, Appli
99	65.8	235 4 US-09-134-001C-4466	Sequence 4466, Ap
100	65.8	333 4 US-09-562-737-4	Sequence 4, Appli

101	25	65.8	367	2	US-08-896-120-1	Sequence 1, Appli	174	25	65.8	1196	4	US-08-881-706-2	Sequence 2, Appli
102	25	65.8	367	2	US-08-896-120-3	Sequence 3, Appli	175	25	65.8	1263	4	US-09-446-504-6	Sequence 6, Appli
103	25	65.8	368	2	US-08-896-120-4	Sequence 4, Appli	176	25	65.8	1263	4	US-09-712-266-6	Sequence 6, Appli
104	25	65.8	393	4	US-09-242-859A-13	Sequence 13, Appli	177	25	65.8	1263	4	US-09-091-889A-4	Sequence 4, Appli
105	25	65.8	409	4	US-09-063-872-2	Sequence 2, Appli	178	25	65.8	1302	4	US-09-423-890-2	Sequence 2, Appli
106	25	65.8	409	4	US-09-667-570A-2	Sequence 2, Appli	179	25	65.8	1327	4	US-09-196-387-2	Sequence 2, Appli
107	25	65.8	410	4	US-09-065-872-1	Sequence 1, Appli	180	25	65.8	1493	4	US-09-423-890-8	Sequence 8, Appli
108	25	65.8	410	4	US-09-667-570A-1	Sequence 1, Appli	181	25	65.8	1593	4	US-08-628-829-4	Sequence 4, Appli
109	25	65.8	419	2	US-08-295-411-1	Sequence 1, Appli	182	25	65.8	2509	1	US-08-469-005A-10	Sequence 10, Appli
110	25	65.8	419	2	US-08-655-471-1	Sequence 1, Appli	183	25	65.8	2511	4	US-09-261-997-2	Sequence 2, Appli
111	25	65.8	419	4	US-09-667-570A-3	Sequence 3, Appli	184	25	65.8	2629	2	US-08-751-189-4	Sequence 4, Appli
112	25	65.8	419	5	PCT-US92-10242-1	Sequence 1, Appli	185	25	65.8	2629	2	US-09-060-862-4	Sequence 4, Appli
113	25	65.8	449	5	US-09-134-10242-5673	Sequence 5673, Ap	186	25	65.8	2629	2	US-09-184-445-4	Sequence 4, Appli
114	25	65.8	460	2	US-08-756-506-2	Sequence 2, Appli	187	25	65.8	2987	2	US-08-970-269A-29	Sequence 29, Appli
115	25	65.8	460	2	US-08-756-506-4	Sequence 4, Appli	188	25	65.8	2987	2	US-09-407-562-29	Sequence 29, Appli
116	25	65.8	460	6	5270178-13	Patent No. 5270178	189	25	65.8	3287	2	US-08-477-451-7	Sequence 7, Appli
117	25	65.8	460	6	5270178-14	Patent No. 5270178	190	25	65.8	3491	2	US-07-642-734C-2	Sequence 2, Appli
118	25	65.8	460	6	5270178-15	Patent No. 5270178	191	25	65.8	3491	3	US-08-439-009A-2	Sequence 2, Appli
119	25	65.8	461	6	5270178-16	Patent No. 5270178	192	25	65.8	3959	2	US-08-970-269A-30	Sequence 30, Appli
120	25	65.8	461	6	5225537-2	Patent No. 5225537	193	25	65.8	3959	2	US-09-407-562-30	Sequence 30, Appli
121	25	65.8	461	6	5270178-17	Patent No. 5270178	194	25	65.8	22	3	US-08-940-095-66	Sequence 66, Appli
122	25	65.8	461	6	5270178-18	Patent No. 5270178	195	25	65.8	22	3	US-08-940-095-66	Sequence 66, Appli
123	25	65.8	461	6	5270178-18	Patent No. 5270178	196	25	65.8	22	3	US-08-940-093-66	Sequence 66, Appli
124	25	65.8	461	6	5460953-3	Patent No. 5460953	197	25	65.8	22	3	US-08-940-093-66	Sequence 66, Appli
125	25	65.8	474	3	US-09-022-699-2	Sequence 2, Appli	198	25	65.8	22	3	US-08-940-096-66	Sequence 66, Appli
126	25	65.8	483	3	US-08-770-544-8	Sequence 8, Appli	199	25	65.8	22	3	US-09-465-719-66	Sequence 66, Appli
127	25	65.8	488	1	US-08-243-542-1	Sequence 1, Appli	200	25	65.8	22	4	US-09-465-719-66	Sequence 66, Appli
128	25	65.8	488	1	US-08-477-407-1	Sequence 1, Appli	201	25	65.8	22	4	US-09-465-719-66	Sequence 66, Appli
129	25	65.8	488	1	US-08-484-355-1	Sequence 1, Appli	202	25	65.8	22	4	US-09-453-605-66	Sequence 66, Appli
130	25	65.8	508	2	US-08-861-464-10	Sequence 10, Appli	203	25	65.8	22	4	US-09-453-605-66	Sequence 66, Appli
131	25	65.8	508	2	US-08-96-001-1	Sequence 10, Appli	204	25	65.8	22	4	US-09-453-605-66	Sequence 66, Appli
132	25	65.8	508	2	US-09-323-433A-10	Sequence 10, Appli	205	25	65.8	22	4	US-09-453-605-66	Sequence 66, Appli
133	25	65.8	524	1	US-08-243-542-2	Sequence 2, Appli	206	25	65.8	33	4	US-09-433-428D-65	Sequence 65, Appli
134	25	65.8	524	1	US-08-477-407-2	Sequence 2, Appli	207	25	65.8	54	2	US-08-977-554-4	Sequence 54, Appli
135	25	65.8	524	1	US-08-484-355-2	Sequence 2, Appli	208	25	65.8	54	4	US-09-225-967-4	Sequence 4, Appli
136	25	65.8	524	3	US-08-556-419-22	Sequence 22, Appli	209	25	65.8	86	4	US-09-227-806-4	Sequence 4, Appli
137	25	65.8	612	3	US-09-212-971-14	Sequence 14, Appli	210	25	65.8	86	4	US-08-936-165A-294	Sequence 294, Appli
138	25	65.8	612	4	US-08-800-929A-14	Sequence 14, Appli	211	25	65.8	119	4	US-09-187-789-21	Sequence 21, Appli
139	25	65.8	612	4	US-08-569-749-14	Sequence 14, Appli	212	25	65.8	119	4	US-09-139-600-16	Sequence 16, Appli
140	25	65.8	612	4	US-09-617-053A-14	Sequence 14, Appli	213	25	65.8	132	4	US-09-134-001C-5049	Sequence 5049, Appli
141	25	65.8	612	5	PCT-US96-12860-14	Sequence 14, Appli	214	25	65.8	142	2	US-08-187-186A-2	Sequence 2, Appli
142	25	65.8	629	3	US-08-556-419-23	Sequence 23, Appli	215	25	65.8	142	2	US-08-442-497C-2	Sequence 2, Appli
143	25	65.8	650	4	US-09-443-793-1	Sequence 1, Appli	216	25	65.8	142	2	US-09-333-033-2	Sequence 2, Appli
144	25	65.8	659	4	US-09-562-737-12	Sequence 12, Appli	217	25	65.8	142	5	PCT-US94-05186-2	Sequence 2, Appli
145	25	65.8	670	1	US-08-477-407-3	Sequence 3, Appli	218	25	65.8	148	4	US-09-786-023-4	Sequence 4, Appli
146	25	65.8	670	1	US-08-484-355-3	Sequence 3, Appli	219	25	65.8	149	4	US-09-433-428D-67	Sequence 67, Appli
147	25	65.8	672	1	US-08-049-254-2	Sequence 2, Appli	220	25	65.8	180	1	US-08-447-010-6	Sequence 6, Appli
148	25	65.8	672	1	US-08-472-934-2	Sequence 2, Appli	221	25	65.8	180	1	US-08-447-010-7	Sequence 7, Appli
149	25	65.8	672	2	US-08-323-460A-2	Sequence 2, Appli	222	25	65.8	183	1	US-08-447-010-5	Sequence 5, Appli
150	25	65.8	672	2	US-08-461-146C-2	Sequence 2, Appli	223	25	65.8	220	4	US-09-433-428D-66	Sequence 66, Appli
151	25	65.8	672	3	US-08-461-145C-2	Sequence 2, Appli	224	25	65.8	224	2	US-08-272-255-16	Sequence 16, Appli
152	25	65.8	672	3	US-08-628-829-2	Sequence 2, Appli	225	25	65.8	224	5	PCT-US95-08565-16	Sequence 16, Appli
153	25	65.8	673	4	US-09-196-387-8	Sequence 8, Appli	226	25	65.8	226	4	US-08-615-192A-327	Sequence 327, Appli
154	25	65.8	683	4	US-08-477-451-42	Sequence 42, Appli	227	25	65.8	237	4	US-08-961-083-136	Sequence 136, Appli
155	25	65.8	707	2	US-09-134-001C-2962	Sequence 2962, Ap	228	25	65.8	256	4	US-09-134-001C-4544	Sequence 4544, Ap
156	25	65.8	708	4	US-09-413-814-18	Sequence 18, Appli	229	25	65.8	271	4	US-09-085-305-12	Sequence 12, Appli
157	25	65.8	728	4	US-09-134-001C-4968	Sequence 4968, Ap	230	25	65.8	307	2	US-09-134-001C-5612	Sequence 5612, Ap
158	25	65.8	734	4	US-09-442-055-2	Sequence 2, Appli	231	25	65.8	306	5	PCT-US95-06119-13	Sequence 13, Appli
159	25	65.8	734	4	US-09-442-055-4	Sequence 4, Appli	232	25	65.8	306	5	US-09-009-895-2	Sequence 2, Appli
160	25	65.8	752	1	US-08-309-512-8	Sequence 8, Appli	233	25	65.8	320	3	US-08-410-167A-4	Sequence 4, Appli
161	25	65.8	752	1	PCT-US92-08756A-8	Sequence 8, Appli	234	25	65.8	330	2	US-08-898-560-1	Sequence 1, Appli
162	25	65.8	769	1	US-08-243-542-4	Sequence 4, Appli	235	25	65.8	330	2	US-09-101-126-1	Sequence 1, Appli
163	25	65.8	769	1	US-08-477-407-4	Sequence 4, Appli	236	25	65.8	343	1	US-08-279-0588-9	Sequence 9, Appli
164	25	65.8	769	1	US-08-484-355-4	Sequence 4, Appli	237	25	65.8	343	1	US-08-828-323-9	Sequence 9, Appli
165	25	65.8	848	4	US-09-540-824-27	Sequence 27, Appli	238	25	65.8	343	1	US-08-153-848-28	Sequence 28, Appli
166	25	65.8	848	4	PCT-US95-13659-4	Sequence 4, Appli	239	25	65.8	355	1	US-08-153-848-32	Sequence 32, Appli
167	25	65.8	887	5	US-09-196-387-10	Sequence 10, Appli	240	25	65.8	355	3	US-09-299-843A-32	Sequence 32, Appli
168	25	65.8	949	2	US-08-956-242-13	Sequence 13, Appli	241	25	65.8	355	3	US-09-306-595C-10	Sequence 10, Appli
169	25	65.8	1159	3	US-09-351-215-13	Sequence 13, Appli	242	25	65.8	355	4	US-09-088-337B-28	Sequence 28, Appli
170	25	65.8	1159	3	US-09-226-012-2	Sequence 2, Appli	243	25	65.8	355	4	US-09-088-337B-32	Sequence 32, Appli
171	25	65.8	1159	4	US-09-226-012-4	Sequence 4, Appli	244	25	65.8	355	5	PCT-US93-11153-28	Sequence 28, Appli
172	25	65.8	1159	4			245	25	65.8				
173	25	65.8	1159	4			246	25	65.8				

247	24	63.2	355	5	PCT-US93-11153-32	Sequence 32, Appl	320	24	63.2	513	4	US-09-488-364-8	Sequence 8, Appl
248	24	63.2	362	2	US-09-080-897-6	Sequence 6, Appl	321	24	63.2	550	4	US-09-167-299-3	Sequence 3, Appl
249	24	63.2	362	4	US-09-323-735-6	Sequence 6, Appl	322	24	63.2	550	1	US-08-121-057-4	Sequence 4, Appl
250	24	63.2	367	4	US-09-134-001C-4168	Sequence 4168, Ap	323	24	63.2	550	2	US-08-509-1870-4	Sequence 4, Appl
251	24	63.2	368	4	US-09-433-428D-58	Sequence 58, Appl	324	24	63.2	550	2	US-09-121-396-4	Sequence 4, Appl
252	24	63.2	374	1	US-08-464-148-2	Sequence 2, Appl	325	24	63.2	579	5	PCT-US93-09704A-4	Sequence 4, Appl
253	24	63.2	374	1	US-08-385-500-2	Sequence 2, Appl	326	24	63.2	579	1	US-08-448-196A-8	Sequence 8, Appl
254	24	63.2	374	1	US-08-846-784-2	Sequence 2, Appl	327	24	63.2	610	4	US-09-455-777-2	Sequence 2, Appl
255	24	63.2	380	1	US-08-420-235B-5	Sequence 5, Appl	328	24	63.2	631	4	US-09-134-001C-4605	Sequence 4605, Ap
256	24	63.2	380	1	US-08-793-624-5	Sequence 5, Appl	329	24	63.2	641	4	US-09-422-869-26	Sequence 26, Appl
257	24	63.2	380	5	PCT-US95-10194-5	Sequence 5, Appl	330	24	63.2	659	4	US-09-562-737-20	Sequence 20, Appl
258	24	63.2	389	2	US-08-846-762-82	Sequence 82, Appl	331	24	63.2	699	4	US-09-134-001C-4054	Sequence 4054, Ap
259	24	63.2	393	2	US-08-977-554-2	Sequence 2, Appl	332	24	63.2	701	4	US-08-923-511-2	Sequence 2, Appl
260	24	63.2	393	4	US-09-225-967-2	Sequence 2, Appl	333	24	63.2	701	4	US-09-422-869-22	Sequence 22, Appl
261	24	63.2	393	4	US-09-227-806-2	Sequence 2, Appl	334	24	63.2	714	4	US-09-416-874A-2	Sequence 2, Appl
262	24	63.2	402	2	US-08-394-189B-14	Sequence 14, Appl	335	24	63.2	740	1	US-08-309-512-10	Sequence 10, Appl
263	24	63.2	402	3	US-08-368-704C-38	Sequence 38, Appl	336	24	63.2	740	1	US-08-309-512-10	Sequence 10, Appl
264	24	63.2	402	3	US-08-954-536-16	Sequence 16, Appl	337	24	63.2	740	1	US-08-309-512-10	Sequence 10, Appl
265	24	63.2	402	4	US-08-908-436-4	Sequence 4, Appl	338	24	63.2	765	5	PCT-US92-08756A-10	Sequence 5, Appl
266	24	63.2	402	5	PCT-US93-05705-14	Sequence 14, Appl	339	24	63.2	851	4	US-09-071-035-326	Sequence 326, App
267	24	63.2	439	4	US-09-433-428D-57	Sequence 57, Appl	340	24	63.2	851	4	US-09-071-035-330	Sequence 330, App
268	24	63.2	460	1	US-08-289-709-1	Sequence 1, Appl	341	24	63.2	890	4	US-09-071-035-334	Sequence 334, App
269	24	63.2	460	1	US-08-602-656-1	Sequence 1, Appl	342	24	63.2	905	4	US-09-342-648-10	Sequence 10, Appl
270	24	63.2	472	4	US-08-914-375C-59	Sequence 59, Appl	343	24	63.2	905	4	US-09-134-001C-3782	Sequence 3782, Ap
271	24	63.2	477	3	US-08-843-982B-8	Sequence 8, Appl	344	24	63.2	912	4	US-08-943-768-2	Sequence 2, Appl
272	24	63.2	478	4	US-09-272-114-1	Sequence 1, Appl	345	24	63.2	912	4	US-09-283-763-2	Sequence 2, Appl
273	24	63.2	478	4	US-09-249-338-2	Sequence 2, Appl	346	24	63.2	930	3	US-09-574-912-2	Sequence 2, Appl
274	24	63.2	489	1	US-08-124-674-2	Sequence 2, Appl	347	24	63.2	972	3	US-08-750-141A-2	Sequence 2, Appl
275	24	63.2	489	1	US-08-589-893-2	Sequence 2, Appl	348	24	63.2	1009	4	US-09-693-146-4	Sequence 4, Appl
276	24	63.2	489	1	US-08-589-893-4	Sequence 4, Appl	349	24	63.2	1041	4	US-09-134-001C-2949	Sequence 2949, Ap
277	24	63.2	489	1	US-08-589-893-6	Sequence 6, Appl	350	24	63.2	1073	4	US-09-541-782-6	Sequence 6, Appl
278	24	63.2	489	1	US-08-589-893-8	Sequence 8, Appl	351	24	63.2	1073	4	US-09-723-820-6	Sequence 6, Appl
279	24	63.2	489	1	US-08-589-893-10	Sequence 10, Appl	352	24	63.2	1128	1	US-08-111-939-2	Sequence 2, Appl
280	24	63.2	489	1	US-08-589-893-12	Sequence 12, Appl	353	24	63.2	1128	4	US-09-641-741-10	Sequence 30, Appl
281	24	63.2	489	1	US-08-589-893-14	Sequence 14, Appl	354	24	63.2	1128	4	US-09-060-445-4	Sequence 4, Appl
282	24	63.2	489	1	US-08-589-893-16	Sequence 16, Appl	355	24	63.2	1130	3	US-08-836-134-23	Sequence 23, Appl
283	24	63.2	489	1	US-08-589-893-18	Sequence 18, Appl	356	24	63.2	1151	4	US-09-493-784-33	Sequence 2, Appl
284	24	63.2	489	1	US-08-589-893-20	Sequence 20, Appl	357	24	63.2	1158	4	US-09-060-445-2	Sequence 2, Appl
285	24	63.2	489	1	US-08-589-893-22	Sequence 22, Appl	358	24	63.2	1232	3	US-09-493-784-2	Sequence 2, Appl
286	24	63.2	489	1	US-08-589-893-24	Sequence 24, Appl	359	24	63.2	1232	3	US-09-493-784-2	Sequence 2, Appl
287	24	63.2	489	2	US-09-020-991-2	Sequence 2, Appl	360	24	63.2	1294	4	US-09-473-717-2	Sequence 2, Appl
288	24	63.2	489	2	US-09-020-991-4	Sequence 4, Appl	361	24	63.2	1305	3	US-08-864-785-3	Sequence 3, Appl
289	24	63.2	489	2	US-09-020-991-6	Sequence 6, Appl	362	24	63.2	1305	3	US-08-894-173-2	Sequence 2, Appl
290	24	63.2	489	2	US-09-020-991-8	Sequence 8, Appl	363	24	63.2	1353	4	US-09-398-193-2	Sequence 2, Appl
291	24	63.2	489	2	US-09-020-991-10	Sequence 10, Appl	364	24	63.2	1353	4	US-09-398-193-2	Sequence 2, Appl
292	24	63.2	489	2	US-09-020-991-12	Sequence 12, Appl	365	24	63.2	1353	4	US-09-473-717-3	Sequence 3, Appl
293	24	63.2	489	2	US-09-020-991-14	Sequence 14, Appl	366	24	63.2	1362	4	US-09-134-001C-4955	Sequence 4955, Ap
294	24	63.2	489	2	US-09-020-991-16	Sequence 16, Appl	367	24	63.2	1462	3	US-07-792-600-31	Sequence 31, Appl
295	24	63.2	489	2	US-09-020-991-18	Sequence 18, Appl	368	24	63.2	1462	3	US-09-157-021-31	Sequence 31, Appl
296	24	63.2	489	2	US-09-020-991-20	Sequence 20, Appl	369	24	63.2	1462	3	US-09-156-842-31	Sequence 31, Appl
297	24	63.2	489	2	US-09-020-991-22	Sequence 22, Appl	370	24	63.2	2628	4	US-09-413-842-11	Sequence 11, Appl
298	24	63.2	489	2	US-09-020-991-24	Sequence 24, Appl	371	24	63.2	2930	4	US-09-417-822-2	Sequence 2, Appl
299	24	63.2	489	2	US-09-062-890-2	Sequence 2, Appl	372	24	63.2	3169	4	US-09-453-702B-257	Sequence 257, App
300	24	63.2	489	2	US-09-062-890-4	Sequence 4, Appl	373	24	63.2	3170	2	US-07-642-734C-5	Sequence 5, Appl
301	24	63.2	489	2	US-09-062-890-6	Sequence 6, Appl	374	24	63.2	3170	2	US-08-439-009A-5	Sequence 5, Appl
302	24	63.2	489	2	US-09-062-890-8	Sequence 8, Appl	375	24	63.2	3170	2	US-08-940-093-104	Sequence 104, App
303	24	63.2	489	2	US-09-062-890-10	Sequence 10, Appl	376	24	63.2	3170	2	US-08-940-093-104	Sequence 104, App
304	24	63.2	489	2	US-09-062-890-12	Sequence 12, Appl	377	24	63.2	3170	2	US-08-940-093-104	Sequence 104, App
305	24	63.2	489	2	US-09-062-890-14	Sequence 14, Appl	378	24	63.2	3170	2	US-08-940-093-104	Sequence 104, App
306	24	63.2	489	2	US-09-062-890-16	Sequence 16, Appl	379	24	63.2	3170	2	US-08-940-093-104	Sequence 104, App
307	24	63.2	489	2	US-09-062-890-18	Sequence 18, Appl	380	24	63.2	3170	2	US-08-940-093-104	Sequence 104, App
308	24	63.2	489	2	US-09-062-890-20	Sequence 20, Appl	381	24	63.2	3170	2	US-08-940-093-104	Sequence 104, App
309	24	63.2	489	2	US-09-062-890-22	Sequence 22, Appl	382	24	63.2	3170	2	US-08-940-093-104	Sequence 104, App
310	24	63.2	489	2	US-09-062-890-24	Sequence 24, Appl	383	24	63.2	3170	2	US-08-940-093-104	Sequence 104, App
311	24	63.2	489	2	US-09-062-890-26	Sequence 26, Appl	384	24	63.2	3170	2	US-08-940-093-104	Sequence 104, App
312	24	63.2	489	2	US-09-062-890-28	Sequence 28, Appl	385	24	63.2	3170	2	US-08-940-093-104	Sequence 104, App
313	24	63.2	489	2	US-09-062-890-30	Sequence 30, Appl	386	24	63.2	3170	2	US-08-940-093-104	Sequence 104, App
314	24	63.2	489	2	US-09-062-890-32	Sequence 32, Appl	387	24	63.2	3170	2	US-08-940-093-104	Sequence 104, App
315	24	63.2	489	2	US-09-062-890-34	Sequence 34, Appl	388	24	63.2	3170	2	US-08-940-093-104	Sequence 104, App
316	24	63.2	489	2	US-09-062-890-36	Sequence 36, Appl	389	24	63.2	3170	2	US-08-940-093-104	Sequence 104, App
317	24	63.2	489	2	US-09-062-890-38	Sequence 38, Appl	390	24	63.2	3170	2	US-08-940-093-104	Sequence 104, App
318	24	63.2	492	4	US-08-462-794-11	Sequence 11, Appl	391	24	63.2	3170	2	US-08-940-093-104	Sequence 104, App
319	24	63.2	513	4	US-08-924-183-8	Sequence 8, Appl	392	24	63.2	3170	2	US-08-940-093-104	Sequence 104, App

393	23	60.5	19	4	US-09-453-838-55	Sequence 55, Appl	466	23	60.5	22	3	US-08-940-093-179	Sequence 179, App
394	23	60.5	19	4	US-09-453-838-60	Sequence 60, Appl	467	23	60.5	22	3	US-08-940-096-4	Sequence 4, Appl
395	23	60.5	21	3	US-08-940-095-67	Sequence 67, Appl	468	23	60.5	22	3	US-08-940-096-7	Sequence 7, Appl
396	23	60.5	21	3	US-08-940-095-85	Sequence 85, Appl	469	23	60.5	22	3	US-08-940-096-9	Sequence 9, Appl
397	23	60.5	21	3	US-08-940-093-67	Sequence 67, Appl	470	23	60.5	22	3	US-08-940-096-10	Sequence 10, Appl
398	23	60.5	21	3	US-08-940-093-85	Sequence 85, Appl	471	23	60.5	22	3	US-08-940-096-12	Sequence 12, Appl
399	23	60.5	21	3	US-08-940-096-67	Sequence 67, Appl	472	23	60.5	22	3	US-08-940-096-16	Sequence 16, Appl
400	23	60.5	21	3	US-08-940-096-85	Sequence 85, Appl	473	23	60.5	22	3	US-08-940-096-18	Sequence 18, Appl
401	23	60.5	21	4	US-09-465-719-67	Sequence 67, Appl	474	23	60.5	22	3	US-08-940-096-21	Sequence 21, Appl
402	23	60.5	21	4	US-09-465-719-85	Sequence 85, Appl	475	23	60.5	22	3	US-08-940-096-24	Sequence 24, Appl
403	23	60.5	21	4	US-09-453-605-67	Sequence 67, Appl	476	23	60.5	22	3	US-08-940-096-29	Sequence 29, Appl
404	23	60.5	21	4	US-09-453-605-85	Sequence 85, Appl	477	23	60.5	22	3	US-08-940-096-30	Sequence 30, Appl
405	23	60.5	21	4	US-09-453-838-67	Sequence 85, Appl	478	23	60.5	22	3	US-08-940-096-32	Sequence 32, Appl
406	23	60.5	21	4	US-09-453-838-85	Sequence 85, Appl	479	23	60.5	22	3	US-08-940-096-34	Sequence 34, Appl
407	23	60.5	22	3	US-08-940-095-4	Sequence 4, Appl	480	23	60.5	22	3	US-08-940-096-35	Sequence 35, Appl
408	23	60.5	22	3	US-08-940-095-7	Sequence 7, Appl	481	23	60.5	22	3	US-08-940-096-48	Sequence 48, Appl
409	23	60.5	22	3	US-08-940-095-9	Sequence 9, Appl	482	23	60.5	22	3	US-08-940-096-53	Sequence 53, Appl
410	23	60.5	22	3	US-08-940-095-10	Sequence 10, Appl	483	23	60.5	22	3	US-08-940-096-56	Sequence 56, Appl
411	23	60.5	22	3	US-08-940-095-12	Sequence 12, Appl	484	23	60.5	22	3	US-08-940-096-59	Sequence 59, Appl
412	23	60.5	22	3	US-08-940-095-14	Sequence 14, Appl	485	23	60.5	22	3	US-08-940-096-79	Sequence 79, Appl
413	23	60.5	22	3	US-08-940-095-16	Sequence 16, Appl	486	23	60.5	22	3	US-08-940-096-80	Sequence 80, Appl
414	23	60.5	22	3	US-08-940-095-18	Sequence 18, Appl	487	23	60.5	22	3	US-08-940-096-86	Sequence 86, Appl
415	23	60.5	22	3	US-08-940-095-21	Sequence 21, Appl	488	23	60.5	22	3	US-08-940-096-100	Sequence 100, App
416	23	60.5	22	3	US-08-940-095-24	Sequence 24, Appl	489	23	60.5	22	3	US-08-940-096-123	Sequence 123, App
417	23	60.5	22	3	US-08-940-095-29	Sequence 29, Appl	490	23	60.5	22	3	US-08-940-096-124	Sequence 124, App
418	23	60.5	22	3	US-08-940-095-30	Sequence 30, Appl	491	23	60.5	22	3	US-08-940-096-125	Sequence 125, App
419	23	60.5	22	3	US-08-940-095-32	Sequence 32, Appl	492	23	60.5	22	3	US-08-940-096-132	Sequence 132, App
420	23	60.5	22	3	US-08-940-095-34	Sequence 34, Appl	493	23	60.5	22	3	US-08-940-096-135	Sequence 135, App
421	23	60.5	22	3	US-08-940-095-35	Sequence 35, Appl	494	23	60.5	22	3	US-08-940-096-178	Sequence 178, App
422	23	60.5	22	3	US-08-940-095-48	Sequence 48, Appl	495	23	60.5	22	3	US-08-940-096-179	Sequence 179, App
423	23	60.5	22	3	US-08-940-095-53	Sequence 53, Appl	496	23	60.5	22	3	US-08-940-096-178	Sequence 178, App
424	23	60.5	22	3	US-08-940-095-56	Sequence 56, Appl	497	23	60.5	22	4	US-08-940-096-179	Sequence 179, App
425	23	60.5	22	3	US-08-940-095-59	Sequence 59, Appl	498	23	60.5	22	4	US-09-329-350-24	Sequence 4, Appl
426	23	60.5	22	3	US-08-940-095-79	Sequence 79, Appl	499	23	60.5	22	4	US-09-465-719-4	Sequence 7, Appl
427	23	60.5	22	3	US-08-940-095-80	Sequence 80, Appl	500	23	60.5	22	4	US-09-465-719-9	Sequence 9, Appl
428	23	60.5	22	3	US-08-940-095-86	Sequence 86, Appl	501	23	60.5	22	4	US-09-465-719-10	Sequence 10, Appl
429	23	60.5	22	3	US-08-940-095-100	Sequence 100, App	502	23	60.5	22	4	US-09-465-719-12	Sequence 12, Appl
430	23	60.5	22	3	US-08-940-095-123	Sequence 123, App	503	23	60.5	22	4	US-09-465-719-14	Sequence 14, Appl
431	23	60.5	22	3	US-08-940-095-124	Sequence 124, App	504	23	60.5	22	4	US-09-465-719-16	Sequence 16, Appl
432	23	60.5	22	3	US-08-940-095-125	Sequence 125, App	505	23	60.5	22	4	US-09-465-719-18	Sequence 18, Appl
433	23	60.5	22	3	US-08-940-095-132	Sequence 132, App	506	23	60.5	22	4	US-09-465-719-21	Sequence 21, Appl
434	23	60.5	22	3	US-08-940-095-135	Sequence 135, App	507	23	60.5	22	4	US-09-465-719-24	Sequence 24, Appl
435	23	60.5	22	3	US-08-940-095-178	Sequence 178, App	508	23	60.5	22	4	US-09-465-719-29	Sequence 29, Appl
436	23	60.5	22	3	US-08-940-095-179	Sequence 179, App	509	23	60.5	22	4	US-09-465-719-30	Sequence 30, Appl
437	23	60.5	22	3	US-08-940-093-4	Sequence 4, Appl	510	23	60.5	22	4	US-09-465-719-32	Sequence 32, Appl
438	23	60.5	22	3	US-08-940-093-7	Sequence 7, Appl	511	23	60.5	22	4	US-09-465-719-34	Sequence 34, Appl
439	23	60.5	22	3	US-08-940-093-9	Sequence 9, Appl	512	23	60.5	22	4	US-09-465-719-48	Sequence 48, Appl
440	23	60.5	22	3	US-08-940-093-10	Sequence 10, Appl	513	23	60.5	22	4	US-09-465-719-53	Sequence 53, Appl
441	23	60.5	22	3	US-08-940-093-12	Sequence 12, Appl	514	23	60.5	22	4	US-09-465-719-56	Sequence 56, Appl
442	23	60.5	22	3	US-08-940-093-14	Sequence 14, Appl	515	23	60.5	22	4	US-09-465-719-59	Sequence 59, Appl
443	23	60.5	22	3	US-08-940-093-16	Sequence 16, Appl	516	23	60.5	22	4	US-09-465-719-79	Sequence 79, Appl
444	23	60.5	22	3	US-08-940-093-18	Sequence 18, Appl	517	23	60.5	22	4	US-09-465-719-80	Sequence 80, Appl
445	23	60.5	22	3	US-08-940-093-21	Sequence 21, Appl	518	23	60.5	22	4	US-09-465-719-86	Sequence 86, Appl
446	23	60.5	22	3	US-08-940-093-24	Sequence 24, Appl	519	23	60.5	22	4	US-09-465-719-100	Sequence 100, App
447	23	60.5	22	3	US-08-940-093-29	Sequence 29, Appl	520	23	60.5	22	4	US-09-465-719-123	Sequence 123, App
448	23	60.5	22	3	US-08-940-093-30	Sequence 30, Appl	521	23	60.5	22	4	US-09-465-719-124	Sequence 124, App
449	23	60.5	22	3	US-08-940-093-32	Sequence 32, Appl	522	23	60.5	22	4	US-09-465-719-125	Sequence 125, App
450	23	60.5	22	3	US-08-940-093-34	Sequence 34, Appl	523	23	60.5	22	4	US-09-465-719-132	Sequence 132, App
451	23	60.5	22	3	US-08-940-093-35	Sequence 35, Appl	524	23	60.5	22	4	US-09-465-719-135	Sequence 135, App
452	23	60.5	22	3	US-08-940-093-48	Sequence 48, Appl	525	23	60.5	22	4	US-09-465-719-178	Sequence 178, App
453	23	60.5	22	3	US-08-940-093-53	Sequence 53, Appl	526	23	60.5	22	4	US-09-465-719-179	Sequence 179, App
454	23	60.5	22	3	US-08-940-093-56	Sequence 56, Appl	527	23	60.5	22	4	US-08-942-597-1	Sequence 1, Appl
455	23	60.5	22	3	US-08-940-093-59	Sequence 59, Appl	528	23	60.5	22	4	US-09-453-605-4	Sequence 4, Appl
456	23	60.5	22	3	US-08-940-093-79	Sequence 79, Appl	529	23	60.5	22	4	US-09-453-605-7	Sequence 7, Appl
457	23	60.5	22	3	US-08-940-093-80	Sequence 80, Appl	530	23	60.5	22	4	US-09-453-605-9	Sequence 9, Appl
458	23	60.5	22	3	US-08-940-093-86	Sequence 86, Appl	531	23	60.5	22	4	US-09-453-605-12	Sequence 12, Appl
459	23	60.5	22	3	US-08-940-093-100	Sequence 100, Appl	532	23	60.5	22	4	US-09-453-605-16	Sequence 16, Appl
460	23	60.5	22	3	US-08-940-093-123	Sequence 123, App	533	23	60.5	22	4	US-09-453-605-18	Sequence 18, Appl
461	23	60.5	22	3	US-08-940-093-124	Sequence 124, App	534	23	60.5	22	4	US-09-453-605-21	Sequence 21, Appl
462	23	60.5	22	3	US-08-940-093-125	Sequence 125, App	535	23	60.5	22	4	US-09-453-605-18	Sequence 18, Appl
463	23	60.5	22	3	US-08-940-093-132	Sequence 132, App	536	23	60.5	22	4	US-09-453-605-21	Sequence 21, Appl
464	23	60.5	22	3	US-08-940-093-135	Sequence 135, App	537	23	60.5	22	4	US-09-453-605-21	Sequence 21, Appl
465	23	60.5	22	3	US-08-940-093-178	Sequence 178, App	538	23	60.5	22	4	US-09-453-605-24	Sequence 24, Appl

539	23	60.5	22	4	US-09-453-605-29	Sequence 29, Appl	612	23	60.5	37	3	US-08-919-597-207	Sequence 207, App
540	23	60.5	22	4	US-09-453-605-30	Sequence 30, Appl	613	23	60.5	37	3	US-08-475-668A-207	Sequence 207, App
541	23	60.5	22	4	US-09-453-605-32	Sequence 32, Appl	614	23	60.5	37	3	US-08-485-551A-207	Sequence 207, App
542	23	60.5	22	4	US-09-453-605-34	Sequence 34, Appl	615	23	60.5	37	3	US-08-471-913A-207	Sequence 207, App
543	23	60.5	22	4	US-09-453-605-35	Sequence 35, Appl	616	23	60.5	37	4	US-08-485-264A-207	Sequence 207, App
544	23	60.5	22	4	US-09-453-605-48	Sequence 48, Appl	617	23	60.5	37	4	US-08-474-349A-207	Sequence 207, App
545	23	60.5	22	4	US-09-453-605-53	Sequence 53, Appl	618	23	60.5	45	1	US-08-407-368-2	Sequence 2, Appl
546	23	60.5	22	4	US-09-453-605-56	Sequence 56, Appl	619	23	60.5	49	4	US-09-085-305-22	Sequence 22, Appl
547	23	60.5	22	4	US-09-453-605-59	Sequence 59, Appl	620	23	60.5	59	4	US-09-085-305-21	Sequence 21, Appl
548	23	60.5	22	4	US-09-453-605-79	Sequence 79, Appl	621	23	60.5	80	4	US-09-085-305-23	Sequence 23, Appl
549	23	60.5	22	4	US-09-453-605-80	Sequence 80, Appl	622	23	60.5	86	4	US-09-085-305-27	Sequence 27, Appl
550	23	60.5	22	4	US-09-453-605-86	Sequence 86, Appl	623	23	60.5	90	4	US-09-085-305-24	Sequence 24, Appl
551	23	60.5	22	4	US-09-453-605-100	Sequence 100, Appl	624	23	60.5	94	3	US-09-147-550-99	Sequence 99, Appl
552	23	60.5	22	4	US-09-453-605-123	Sequence 123, Appl	625	23	60.5	94	4	US-09-557-917-99	Sequence 99, Appl
553	23	60.5	22	4	US-09-453-605-124	Sequence 124, Appl	626	23	60.5	99	1	US-08-241-853-14	Sequence 14, Appl
554	23	60.5	22	4	US-09-453-605-125	Sequence 125, Appl	627	23	60.5	99	2	US-08-850-917-14	Sequence 14, Appl
555	23	60.5	22	4	US-09-453-605-132	Sequence 132, Appl	628	23	60.5	101	3	US-09-050-603A-37	Sequence 37, Appl
556	23	60.5	22	4	US-09-453-605-135	Sequence 135, Appl	629	23	60.5	101	3	US-09-102-420B-37	Sequence 37, Appl
557	23	60.5	22	4	US-09-453-605-178	Sequence 178, Appl	630	23	60.5	101	4	US-09-497-698-37	Sequence 37, Appl
558	23	60.5	22	4	US-09-453-605-179	Sequence 179, Appl	631	23	60.5	112	6	5494663-7	Patent No. 5494663
559	23	60.5	22	4	US-09-453-838-4	Sequence 4, Appl	632	23	60.5	118	4	US-09-085-305-2	Sequence 2, Appl
560	23	60.5	22	4	US-09-453-838-7	Sequence 7, Appl	633	23	60.5	119	4	US-09-615-192A-383	Sequence 383, Appl
561	23	60.5	22	4	US-09-453-838-9	Sequence 9, Appl	634	23	60.5	135	1	US-08-426-627-15	Sequence 15, Appl
562	23	60.5	22	4	US-09-453-838-10	Sequence 10, Appl	635	23	60.5	139	4	US-09-085-305-26	Sequence 26, Appl
563	23	60.5	22	4	US-09-453-838-12	Sequence 12, Appl	636	23	60.5	140	4	US-08-961-083-100	Sequence 100, Appl
564	23	60.5	22	4	US-09-453-838-14	Sequence 14, Appl	637	23	60.5	142	4	US-09-072-596-277	Sequence 277, Appl
565	23	60.5	22	4	US-09-453-838-16	Sequence 16, Appl	638	23	60.5	144	4	US-09-370-838-73	Sequence 73, Appl
566	23	60.5	22	4	US-09-453-838-18	Sequence 18, Appl	639	23	60.5	148	2	US-09-193-877-1	Sequence 1, Appl
567	23	60.5	22	4	US-09-453-838-21	Sequence 21, Appl	640	23	60.5	149	4	US-09-085-305-25	Sequence 25, Appl
568	23	60.5	22	4	US-09-453-838-24	Sequence 24, Appl	641	23	60.5	150	4	US-09-325-932A-155	Sequence 155, Appl
569	23	60.5	22	4	US-09-453-838-29	Sequence 29, Appl	642	23	60.5	154	1	US-08-318-193-77	Sequence 77, Appl
570	23	60.5	22	4	US-09-453-838-30	Sequence 30, Appl	643	23	60.5	155	5	US-08-955-848A-1	Sequence 1, Appl
571	23	60.5	22	4	US-09-453-838-32	Sequence 32, Appl	644	23	60.5	155	5	PCT-US95-03866-2	Sequence 2, Appl
572	23	60.5	22	4	US-09-453-838-34	Sequence 34, Appl	645	23	60.5	166	2	US-08-628-428-2	Sequence 2, Appl
573	23	60.5	22	4	US-09-453-838-35	Sequence 35, Appl	646	23	60.5	166	2	US-08-628-428-5	Sequence 5, Appl
574	23	60.5	22	4	US-09-453-838-48	Sequence 48, Appl	647	23	60.5	166	2	US-08-628-428-8	Sequence 8, Appl
575	23	60.5	22	4	US-09-453-838-53	Sequence 53, Appl	648	23	60.5	166	2	US-09-106-891-2	Sequence 2, Appl
576	23	60.5	22	4	US-09-453-838-56	Sequence 56, Appl	649	23	60.5	166	4	US-08-172-507-2	Sequence 2, Appl
577	23	60.5	22	4	US-09-453-838-59	Sequence 59, Appl	650	23	60.5	166	5	PCT-US95-03866-34	Sequence 34, Appl
578	23	60.5	22	4	US-09-453-838-79	Sequence 79, Appl	651	23	60.5	166	5	PCT-US95-03866-36	Sequence 36, Appl
579	23	60.5	22	4	US-09-453-838-80	Sequence 80, Appl	652	23	60.5	166	5	PCT-US95-03866-36	Sequence 36, Appl
580	23	60.5	22	4	US-09-453-838-86	Sequence 86, Appl	653	23	60.5	167	5	PCT-US95-03866-20	Sequence 20, Appl
581	23	60.5	22	4	US-09-453-838-100	Sequence 100, Appl	654	23	60.5	167	2	US-08-955-848A-2	Sequence 2, Appl
582	23	60.5	22	4	US-09-453-838-123	Sequence 123, Appl	655	23	60.5	169	2	US-08-955-848A-3	Sequence 3, Appl
583	23	60.5	22	4	US-09-453-838-124	Sequence 124, Appl	656	23	60.5	169	2	US-08-955-848A-10	Sequence 10, Appl
584	23	60.5	22	4	US-09-453-838-125	Sequence 125, Appl	657	23	60.5	169	2	US-08-955-848A-11	Sequence 11, Appl
585	23	60.5	22	4	US-09-453-838-132	Sequence 132, Appl	658	23	60.5	169	2	US-08-955-848A-12	Sequence 12, Appl
586	23	60.5	22	4	US-09-453-838-135	Sequence 135, Appl	659	23	60.5	169	2	US-08-955-848A-13	Sequence 13, Appl
587	23	60.5	22	4	US-09-453-838-178	Sequence 178, Appl	660	23	60.5	169	2	US-08-955-848A-15	Sequence 15, Appl
588	23	60.5	22	4	US-09-453-838-179	Sequence 179, Appl	661	23	60.5	169	2	US-08-955-848A-16	Sequence 16, Appl
589	23	60.5	22	3	US-08-940-095-2	Sequence 2, Appl	662	23	60.5	169	2	US-08-955-848A-17	Sequence 17, Appl
590	23	60.5	22	3	US-08-940-095-5	Sequence 5, Appl	663	23	60.5	169	2	US-08-955-848A-18	Sequence 18, Appl
591	23	60.5	22	3	US-08-940-095-2	Sequence 2, Appl	664	23	60.5	169	2	US-08-955-848A-11	Sequence 11, Appl
592	23	60.5	22	3	US-08-940-093-2	Sequence 2, Appl	665	23	60.5	169	2	US-08-955-848A-13	Sequence 13, Appl
593	23	60.5	22	3	US-08-940-093-5	Sequence 5, Appl	666	23	60.5	169	2	US-08-955-848A-14	Sequence 14, Appl
594	23	60.5	22	3	US-08-940-096-2	Sequence 2, Appl	667	23	60.5	169	2	US-08-955-848A-15	Sequence 15, Appl
595	23	60.5	22	3	US-08-940-096-5	Sequence 5, Appl	668	23	60.5	169	2	US-08-955-848A-16	Sequence 16, Appl
596	23	60.5	22	3	US-09-465-719-2	Sequence 2, Appl	669	23	60.5	169	2	US-08-955-848A-17	Sequence 17, Appl
597	23	60.5	22	3	US-09-465-719-5	Sequence 5, Appl	670	23	60.5	169	2	US-08-955-848A-18	Sequence 18, Appl
598	23	60.5	22	3	US-09-453-605-2	Sequence 2, Appl	671	23	60.5	169	2	US-08-955-848A-19	Sequence 19, Appl
599	23	60.5	22	3	US-09-453-605-5	Sequence 5, Appl	672	23	60.5	169	2	US-08-955-848A-20	Sequence 20, Appl
600	23	60.5	22	3	US-09-453-838-2	Sequence 2, Appl	673	23	60.5	169	2	US-08-955-848A-21	Sequence 21, Appl
601	23	60.5	22	3	US-09-453-838-5	Sequence 5, Appl	674	23	60.5	169	2	US-08-955-848A-22	Sequence 22, Appl
602	23	60.5	22	3	US-09-082-279B-454	Sequence 454, Appl	675	23	60.5	169	2	US-08-955-848A-23	Sequence 23, Appl
603	23	60.5	22	3	US-09-082-279B-455	Sequence 455, Appl	676	23	60.5	169	2	US-08-955-848A-24	Sequence 24, Appl
604	23	60.5	22	3	US-09-082-279B-456	Sequence 456, Appl	677	23	60.5	169	2	US-08-955-848A-25	Sequence 25, Appl
605	23	60.5	22	3	US-08-474-349A-481	Sequence 481, Appl	678	23	60.5	169	2	US-08-955-848A-26	Sequence 26, Appl
606	23	60.5	22	3	US-08-474-349A-482	Sequence 482, Appl	679	23	60.5	169	2	US-08-955-848A-27	Sequence 27, Appl
607	23	60.5	22	3	US-08-474-349A-483	Sequence 483, Appl	680	23	60.5	169	2	US-08-955-848A-28	Sequence 28, Appl
608	23	60.5	22	3	US-09-315-304B-454	Sequence 454, Appl	681	23	60.5	169	2	US-08-955-848A-29	Sequence 29, Appl
609	23	60.5	22	3	US-09-315-304B-455	Sequence 455, Appl	682	23	60.5	169	2	US-08-955-848A-30	Sequence 30, Appl
610	23	60.5	22	3	US-08-486-099-207	Sequence 207, Appl	683	23	60.5	169	2	US-08-955-848A-31	Sequence 31, Appl
611	23	60.5	22	3	US-08-484-223B-207	Sequence 207, Appl	684	23	60.5	169	2	US-08-955-848A-32	Sequence 32, Appl

685	23	60.5	169	2	US-08-955-848A-33	Sequence 33, Appl	758	23	60.5	273	4	US-08-482-918-48	Sequence 48, Appl
686	23	60.5	169	2	US-08-955-848A-34	Sequence 34, Appl	759	23	60.5	273	4	US-08-482-918-49	Sequence 49, Appl
687	23	60.5	169	2	US-08-955-848A-35	Sequence 35, Appl	760	23	60.5	273	4	US-08-482-918-50	Sequence 50, Appl
688	23	60.5	169	2	US-08-955-848A-36	Sequence 36, Appl	761	23	60.5	273	4	US-08-482-918-51	Sequence 51, Appl
689	23	60.5	174	3	US-08-906-136A-2	Sequence 2, Appl	762	23	60.5	273	4	US-08-482-918-52	Sequence 52, Appl
690	23	60.5	174	3	US-08-906-136A-4	Sequence 4, Appl	763	23	60.5	273	4	US-08-482-918-53	Sequence 53, Appl
691	23	60.5	174	3	US-08-906-136A-2	Sequence 2, Appl	764	23	60.5	273	4	US-08-482-918-54	Sequence 54, Appl
692	23	60.5	174	3	US-08-906-136A-4	Sequence 4, Appl	765	23	60.5	273	4	US-08-482-918-55	Sequence 55, Appl
693	23	60.5	178	4	US-09-363-427-4	Sequence 2, Appl	766	23	60.5	273	4	US-08-482-918-56	Sequence 56, Appl
694	23	60.5	179	2	US-08-993-228-33	Sequence 3654, Ap	767	23	60.5	273	4	US-08-482-918-57	Sequence 57, Appl
695	23	60.5	182	1	US-07-949-812-6	Sequence 33, Appl	768	23	60.5	273	4	US-08-482-918-58	Sequence 58, Appl
696	23	60.5	183	4	US-09-143-211-18	Sequence 18, Appl	769	23	60.5	273	4	US-08-482-918-59	Sequence 59, Appl
697	23	60.5	190	4	US-09-080-662-2	Sequence 2, Appl	770	23	60.5	273	4	US-08-482-918-60	Sequence 60, Appl
698	23	60.5	190	4	US-09-080-662-4	Sequence 4, Appl	771	23	60.5	273	4	US-08-482-918-61	Sequence 61, Appl
699	23	60.5	195	4	US-08-482-918-44	Sequence 44, Appl	772	23	60.5	273	4	US-08-482-918-62	Sequence 62, Appl
700	23	60.5	195	4	US-08-482-918-44	Sequence 44, Appl	773	23	60.5	273	4	US-08-482-918-63	Sequence 63, Appl
701	23	60.5	195	4	US-09-134-001C-3217	Sequence 3217, Ap	774	23	60.5	273	4	US-08-482-918-64	Sequence 64, Appl
702	23	60.5	196	4	US-08-336-728A-44	Sequence 44, Appl	775	23	60.5	273	4	US-08-482-918-65	Sequence 65, Appl
703	23	60.5	197	2	US-09-071-035-488	Sequence 48, Appl	776	23	60.5	273	4	US-08-482-918-66	Sequence 66, Appl
704	23	60.5	199	2	US-08-700-607-1	Sequence 1, Appl	777	23	60.5	273	4	US-08-482-918-67	Sequence 67, Appl
705	23	60.5	201	4	US-09-134-001C-3343	Sequence 3343, Ap	778	23	60.5	273	4	US-08-482-918-68	Sequence 68, Appl
706	23	60.5	205	1	US-08-133-979A-4	Sequence 4, Appl	779	23	60.5	273	4	US-08-482-918-69	Sequence 69, Appl
707	23	60.5	205	2	US-08-436-890-4	Sequence 4, Appl	780	23	60.5	273	4	US-08-482-918-70	Sequence 70, Appl
708	23	60.5	205	2	US-08-436-890-4	Sequence 4, Appl	781	23	60.5	273	4	US-08-482-918-71	Sequence 71, Appl
709	23	60.5	207	4	US-09-143-211-3	Sequence 3, Appl	782	23	60.5	273	4	US-08-482-918-72	Sequence 72, Appl
710	23	60.5	208	4	US-08-836-252A-6	Sequence 6, Appl	783	23	60.5	273	4	US-08-482-918-73	Sequence 73, Appl
711	23	60.5	208	4	US-08-482-918-46	Sequence 46, Appl	784	23	60.5	273	4	US-08-482-918-74	Sequence 74, Appl
712	23	60.5	208	4	US-09-224-681-46	Sequence 46, Appl	785	23	60.5	273	4	US-08-482-918-75	Sequence 75, Appl
713	23	60.5	208	4	US-08-336-728A-46	Sequence 46, Appl	786	23	60.5	273	4	US-08-482-918-76	Sequence 76, Appl
714	23	60.5	212	2	US-08-477-396A-2	Sequence 2, Appl	787	23	60.5	273	4	US-08-482-918-77	Sequence 77, Appl
715	23	60.5	214	2	US-08-984-172-1	Sequence 1, Appl	788	23	60.5	273	4	US-08-482-918-78	Sequence 78, Appl
716	23	60.5	216	4	US-09-134-001C-5630	Sequence 5630, Ap	789	23	60.5	273	4	US-08-482-918-79	Sequence 79, Appl
717	23	60.5	223	4	US-09-134-001C-4407	Sequence 4407, Ap	790	23	60.5	273	4	US-08-482-918-80	Sequence 80, Appl
718	23	60.5	227	4	US-09-230-196-21	Sequence 21, Appl	791	23	60.5	273	4	US-08-482-918-81	Sequence 81, Appl
719	23	60.5	231	4	US-09-134-001C-4397	Sequence 4397, Ap	792	23	60.5	273	4	US-08-482-918-82	Sequence 82, Appl
720	23	60.5	244	4	US-08-486-099-104	Sequence 104, App	793	23	60.5	273	4	US-08-482-918-83	Sequence 83, Appl
721	23	60.5	245	3	US-08-360-107A-114	Sequence 114, App	794	23	60.5	273	4	US-08-482-918-84	Sequence 84, Appl
722	23	60.5	245	3	US-08-484-223B-104	Sequence 104, App	795	23	60.5	273	4	US-08-482-918-85	Sequence 85, Appl
723	23	60.5	245	3	US-08-484-223B-104	Sequence 104, App	796	23	60.5	273	4	US-08-482-918-86	Sequence 86, Appl
724	23	60.5	245	3	US-08-484-223B-104	Sequence 104, App	797	23	60.5	273	4	US-08-482-918-87	Sequence 87, Appl
725	23	60.5	245	3	US-08-484-223B-104	Sequence 104, App	798	23	60.5	273	4	US-08-482-918-88	Sequence 88, Appl
726	23	60.5	245	3	US-08-484-223B-104	Sequence 104, App	799	23	60.5	273	4	US-08-482-918-89	Sequence 89, Appl
727	23	60.5	245	3	US-08-484-223B-104	Sequence 104, App	800	23	60.5	273	4	US-08-482-918-90	Sequence 90, Appl
728	23	60.5	245	3	US-08-484-223B-104	Sequence 104, App	801	23	60.5	273	4	US-08-482-918-91	Sequence 91, Appl
729	23	60.5	245	3	US-08-484-223B-104	Sequence 104, App	802	23	60.5	273	4	US-08-482-918-92	Sequence 92, Appl
730	23	60.5	245	3	US-08-484-223B-104	Sequence 104, App	803	23	60.5	273	4	US-08-482-918-93	Sequence 93, Appl
731	23	60.5	245	3	US-08-484-223B-104	Sequence 104, App	804	23	60.5	273	4	US-08-482-918-94	Sequence 94, Appl
732	23	60.5	245	3	US-08-484-223B-104	Sequence 104, App	805	23	60.5	273	4	US-08-482-918-95	Sequence 95, Appl
733	23	60.5	245	3	US-08-484-223B-104	Sequence 104, App	806	23	60.5	273	4	US-08-482-918-96	Sequence 96, Appl
734	23	60.5	245	3	US-08-484-223B-104	Sequence 104, App	807	23	60.5	273	4	US-08-482-918-97	Sequence 97, Appl
735	23	60.5	245	3	US-08-484-223B-104	Sequence 104, App	808	23	60.5	273	4	US-08-482-918-98	Sequence 98, Appl
736	23	60.5	245	3	US-08-484-223B-104	Sequence 104, App	809	23	60.5	273	4	US-08-482-918-99	Sequence 99, Appl
737	23	60.5	245	3	US-08-484-223B-104	Sequence 104, App	810	23	60.5	273	4	US-08-482-918-100	Sequence 100, Appl
738	23	60.5	245	3	US-08-484-223B-104	Sequence 104, App	811	23	60.5	273	4	US-08-482-918-101	Sequence 101, Appl
739	23	60.5	245	3	US-08-484-223B-104	Sequence 104, App	812	23	60.5	273	4	US-08-482-918-102	Sequence 102, Appl
740	23	60.5	245	3	US-08-484-223B-104	Sequence 104, App	813	23	60.5	273	4	US-08-482-918-103	Sequence 103, Appl
741	23	60.5	245	3	US-08-484-223B-104	Sequence 104, App	814	23	60.5	273	4	US-08-482-918-104	Sequence 104, Appl
742	23	60.5	245	3	US-08-484-223B-104	Sequence 104, App	815	23	60.5	273	4	US-08-482-918-105	Sequence 105, Appl
743	23	60.5	245	3	US-08-484-223B-104	Sequence 104, App	816	23	60.5	273	4	US-08-482-918-106	Sequence 106, Appl
744	23	60.5	245	3	US-08-484-223B-104	Sequence 104, App	817	23	60.5	273	4	US-08-482-918-107	Sequence 107, Appl
745	23	60.5	245	3	US-08-484-223B-104	Sequence 104, App	818	23	60.5	273	4	US-08-482-918-108	Sequence 108, Appl
746	23	60.5	245	3	US-08-484-223B-104	Sequence 104, App	819	23	60.5	273	4	US-08-482-918-109	Sequence 109, Appl
747	23	60.5	245	3	US-08-484-223B-104	Sequence 104, App	820	23	60.5	273	4	US-08-482-918-110	Sequence 110, Appl
748	23	60.5	245	3	US-08-484-223B-104	Sequence 104, App	821	23	60.5	273	4	US-08-482-918-111	Sequence 111, Appl
749	23	60.5	245	3	US-08-484-223B-104	Sequence 104, App	822	23	60.5	273	4	US-08-482-918-112	Sequence 112, Appl
750	23	60.5	245	3	US-08-484-223B-104	Sequence 104, App	823	23	60.5	273	4	US-08-482-918-113	Sequence 113, Appl
751	23	60.5	245	3	US-08-484-223B-104	Sequence 104, App	824	23	60.5	273	4	US-08-482-918-114	Sequence 114, Appl
752	23	60.5	245	3	US-08-484-223B-104	Sequence 104, App	825	23	60.5	273	4	US-08-482-918-115	Sequence 115, Appl
753	23	60.5	245	3	US-08-484-223B-104	Sequence 104, App	826	23	60.5	273	4	US-08-482-918-116	Sequence 116, Appl
754	23	60.5	245	3	US-08-484-223B-104	Sequence 104, App	827	23	60.5	273	4	US-08-482-918-117	Sequence 117, Appl
755	23	60.5	245	3	US-08-484-223B-104	Sequence 104, App	828	23	60.5	273	4	US-08-482-918-118	Sequence 118, Appl
756	23	60.5	245	3	US-08-484-223B-104	Sequence 104, App	829	23	60.5	273	4	US-08-482-918-119	Sequence 119, Appl
757	23	60.5	245	3	US-08-484-223B-104	Sequence 104, App	830	23	60.5	273	4	US-08-482-918-120	Sequence 120, Appl

831	23	60.5	426	4	US-09-134-001C-3346	Sequence 3346, Ap	904	23	60.5	719	4	US-08-854-050-7	Sequence 7, Appli
832	23	60.5	440	2	US-08-808-931-24	Sequence 24, Appl	905	23	60.5	719	4	US-09-430-323-7	Sequence 7, Appli
833	23	60.5	440	3	US-08-808-931-24	Sequence 24, Appl	906	23	60.5	740	4	US-09-624-693A-15	Sequence 15, Appl
834	23	60.5	440	3	US-09-050-603A-24	Sequence 24, Appl	907	23	60.5	777	2	US-08-477-396A-4	Sequence 4, Appli
835	23	60.5	440	3	US-09-102-420B-24	Sequence 24, Appl	908	23	60.5	779	1	US-08-426-627-24	Sequence 24, Appli
836	23	60.5	440	3	US-09-497-698-24	Sequence 24, Appl	909	23	60.5	779	1	US-08-426-627-24	Sequence 24, Appli
837	23	60.5	441	2	US-08-713-815A-3	Sequence 3, Appli	910	23	60.5	790	1	US-08-363-560-2	Sequence 2, Appli
838	23	60.5	459	4	US-09-071-709-1	Sequence 1, Appli	911	23	60.5	796	4	US-08-961-083-56	Sequence 56, Appli
839	23	60.5	481	1	US-08-472-028A-6	Sequence 6, Appli	912	23	60.5	802	4	US-09-433-043B-120	Sequence 120, App
840	23	60.5	481	2	US-08-808-931-6	Sequence 6, Appli	913	23	60.5	814	4	US-09-813-819-4	Sequence 4, Appli
841	23	60.5	481	3	US-08-808-931-6	Sequence 6, Appli	914	23	60.5	814	4	US-09-920-048-4	Sequence 4, Appli
842	23	60.5	481	3	US-09-050-603A-6	Sequence 6, Appli	915	23	60.5	836	1	US-08-426-627-6	Sequence 6, Appli
843	23	60.5	481	3	US-09-102-420B-6	Sequence 6, Appli	916	23	60.5	837	1	US-08-426-627-23	Sequence 23, Appli
844	23	60.5	481	4	US-09-015-683-6	Sequence 6, Appli	917	23	60.5	855	4	US-09-813-819-2	Sequence 2, Appli
845	23	60.5	481	4	US-09-497-698-6	Sequence 6, Appli	918	23	60.5	855	4	US-09-920-048-2	Sequence 2, Appli
846	23	60.5	483	4	US-09-071-296-6	Sequence 6, Appli	919	23	60.5	894	4	US-09-735-934A-4	Sequence 4, Appli
847	23	60.5	483	4	US-09-196-268-6	Sequence 6, Appli	920	23	60.5	908	5	PCT-US95-03747-3	Sequence 3, Appli
848	23	60.5	483	4	US-09-191-998-6	Sequence 6, Appli	921	23	60.5	912	4	US-09-134-001C-3993	Sequence 2993, Ap
849	23	60.5	498	1	US-08-466-855A-6	Sequence 6, Appli	922	23	60.5	936	5	PCT-US94-05905-22	Sequence 22, Appli
850	23	60.5	498	2	US-08-466-859-12	Sequence 12, Appl	923	23	60.5	938	2	US-08-906-488-2	Sequence 2, Appli
851	23	60.5	498	2	US-08-700-636-12	Sequence 12, Appl	924	23	60.5	939	5	PCT-US94-05905-20	Sequence 20, Appli
852	23	60.5	498	3	US-08-467-574-12	Sequence 12, Appl	925	23	60.5	953	4	US-09-245-281-43	Sequence 43, Appli
853	23	60.5	498	4	US-09-217-345-12	Sequence 12, Appl	926	23	60.5	953	4	US-09-207-359B-43	Sequence 43, Appli
854	23	60.5	498	4	US-08-487-596-18	Sequence 18, Appl	927	23	60.5	966	4	US-09-207-359B-47	Sequence 47, Appli
855	23	60.5	501	4	US-09-465-519-2	Sequence 2, Appli	928	23	60.5	967	4	US-09-130-491-2	Sequence 2, Appli
856	23	60.5	501	4	US-09-465-519-4	Sequence 4, Appli	929	23	60.5	983	4	US-09-134-001C-3814	Sequence 3814, Ap
857	23	60.5	502	4	US-09-413-814-106	Sequence 106, App	930	23	60.5	1030	4	US-09-091-117-2	Sequence 2, Appli
858	23	60.5	502	4	US-09-342-647-4	Sequence 4, Appli	931	23	60.5	1075	2	US-08-993-228-19	Sequence 19, Appl
859	23	60.5	506	4	US-09-134-001C-4733	Sequence 4733, Ap	932	23	60.5	1163	4	US-09-178-252-21	Sequence 21, Appli
860	23	60.5	514	4	US-09-370-098-2	Sequence 2, Appli	933	23	60.5	1248	2	US-09-080-897-2	Sequence 2, Appli
861	23	60.5	523	4	US-09-555-889A-2	Sequence 2, Appli	934	23	60.5	1248	4	US-09-323-735-2	Sequence 2, Appli
862	23	60.5	524	4	US-09-230-388-1	Sequence 1, Appli	935	23	60.5	1255	2	US-09-080-897-4	Sequence 4, Appli
863	23	60.5	526	1	US-07-921-796-6	Sequence 6, Appli	936	23	60.5	1255	3	US-08-829-595-1	Sequence 1, Appli
864	23	60.5	526	1	US-07-921-796-8	Sequence 8, Appli	937	23	60.5	1255	4	US-09-323-735-4	Sequence 4, Appli
865	23	60.5	528	1	US-08-368-071-13	Sequence 13, Appl	938	23	60.5	1257	1	US-08-340-428B-49	Sequence 49, Appli
866	23	60.5	528	1	US-08-458-181-13	Sequence 13, Appl	939	23	60.5	1265	4	US-09-347-878-5	Sequence 5, Appli
867	23	60.5	528	5	PCT-US93-02172-13	Sequence 13, Appl	940	23	60.5	1365	4	US-09-347-878-7	Sequence 7, Appli
868	23	60.5	548	4	US-09-738-894A-4	Sequence 4, Appli	941	23	60.5	1313	4	US-09-071-035-450	Sequence 450, App
869	23	60.5	548	4	US-09-398-395A-2	Sequence 2, Appli	942	23	60.5	1313	4	US-09-071-035-454	Sequence 454, App
870	23	60.5	548	4	US-09-398-395A-4	Sequence 4, Appli	943	23	60.5	1315	3	US-08-899-595-3	Sequence 3, Appli
871	23	60.5	548	4	US-09-398-395A-6	Sequence 6, Appli	944	23	60.5	1489	6	5183745-2	Sequence 3, Appli
872	23	60.5	548	4	US-09-398-395A-8	Sequence 8, Appli	945	23	60.5	1529	4	US-09-134-001C-3945	Sequence 3945, Ap
873	23	60.5	548	4	US-09-398-395A-10	Sequence 10, Appl	946	23	60.5	1568	4	US-09-181-706-2	Sequence 2, Appli
874	23	60.5	548	4	US-09-398-395A-12	Sequence 12, Appl	947	23	60.5	1568	4	US-09-458-791-2	Sequence 2, Appli
875	23	60.5	550	2	US-08-443-639-8	Sequence 8, Appli	948	23	60.5	1568	4	US-09-459-066-2	Sequence 2, Appli
876	23	60.5	550	3	US-08-577-483-8	Sequence 8, Appli	949	23	60.5	1705	4	US-08-669-785-4	Sequence 4, Appli
877	23	60.5	553	3	US-08-501-572-2	Sequence 2, Appli	950	23	60.5	1706	4	US-08-669-785-2	Sequence 2, Appli
878	23	60.5	553	3	US-09-040-444-4	Sequence 4, Appli	951	23	60.5	1794	6	5183745-6	Sequence 6, Appli
879	23	60.5	553	4	US-08-997-251-4	Sequence 2, Appli	952	23	60.5	1872	1	US-08-188-582-14	Sequence 14, Appl
880	23	60.5	553	4	US-09-738-894A-2	Sequence 2, Appli	953	23	60.5	1872	1	US-08-646-715-14	Sequence 14, Appl
881	23	60.5	553	4	US-09-802-117-2	Sequence 2, Appli	954	23	60.5	1893	1	US-08-188-582-11	Sequence 11, Appl
882	23	60.5	552	2	US-08-599-171A-30	Sequence 30, Appl	955	23	60.5	1893	1	US-08-646-715-11	Sequence 11, Appl
883	23	60.5	552	2	US-08-646-590B-30	Sequence 30, Appl	956	23	60.5	2037	4	US-09-306-998-3	Sequence 3, Appli
884	23	60.5	592	3	US-09-069-226-30	Sequence 30, Appl	957	23	60.5	2154	2	US-08-841-349-4	Sequence 4, Appli
885	23	60.5	592	3	US-09-412-184-30	Sequence 30, Appl	958	23	60.5	2237	1	US-08-354-973-1	Sequence 1, Appli
886	23	60.5	608	4	US-09-130-491-13	Sequence 13, Appl	959	23	60.5	2237	4	US-09-413-814-70	Sequence 70, Appli
887	23	60.5	613	1	US-08-272-875-3	Sequence 3, Appli	960	23	60.5	2293	4	US-09-368-590-2	Sequence 2, Appli
888	23	60.5	613	1	US-09-298-367B-11	Sequence 11, Appl	961	23	60.5	2391	2	US-08-446-855A-2	Sequence 2, Appli
889	23	60.5	627	4	US-09-071-035-452	Sequence 452, App	962	23	60.5	2391	4	US-09-150-741-2	Sequence 2, Appli
890	23	60.5	630	4	US-09-342-647-2	Sequence 2, Appli	963	23	60.5	2439	3	US-09-335-409-7	Sequence 7, Appli
891	23	60.5	630	4	US-09-562-737-14	Sequence 14, Appl	964	23	60.5	2439	3	US-09-568-102-7	Sequence 7, Appli
892	23	60.5	659	4	US-09-562-737-15	Sequence 15, Appl	965	23	60.5	2439	4	US-09-567-966-7	Sequence 7, Appli
893	23	60.5	664	1	US-08-421-661-6	Sequence 6, Appli	966	23	60.5	2439	4	US-09-568-480-7	Sequence 7, Appli
894	23	60.5	667	4	US-09-071-709-9	Sequence 9, Appli	967	23	60.5	2439	4	US-09-568-486-7	Sequence 7, Appli
895	23	60.5	667	4	US-09-342-647-28	Sequence 28, Appli	968	23	60.5	2439	4	US-09-568-472-7	Sequence 7, Appli
896	23	60.5	684	4	US-09-134-001C-4775	Sequence 4775, Ap	969	23	60.5	2439	4	US-09-567-899-7	Sequence 7, Appli
897	23	60.5	699	3	US-08-851-843A-52	Sequence 52, Appl	970	23	60.5	2627	2	US-08-751-189-3	Sequence 3, Appli
898	23	60.5	699	3	US-08-851-843A-188	Sequence 188, App	971	23	60.5	2627	2	US-09-060-836-3	Sequence 3, Appli
899	23	60.5	699	4	US-08-854-050-52	Sequence 52, Appl	972	23	60.5	2627	5	US-09-184-445-3	Sequence 3, Appli
900	23	60.5	699	4	US-09-430-323-52	Sequence 52, Appl	973	23	60.5	2749	4	PCT-US96-01775-1	Sequence 1, Appli
901	23	60.5	705	4	US-09-134-001C-5356	Sequence 5356, Ap	974	23	60.5	2749	4	US-09-385-222A-4	Sequence 4, Appli
902	23	60.5	719	3	US-08-851-843A-7	Sequence 7, Appli	975	23	60.5	2802	4	US-09-542-331-1	Sequence 1, Appli
903	23	60.5	719	4	US-08-974-549A-219	Sequence 219, App	976	23	60.5	2802	4	US-09-510-791-1	Sequence 1, Appli



```

977 23 60.5 3038 1 US-08-450-332-2 Sequence 2, Appll
978 23 60.5 3038 2 US-08-637-640-2 Sequence 2, Appll
979 23 60.5 3038 3 US-09-004-406C-2 Sequence 2, Appll
980 23 60.5 3567 2 US-07-642-734C-4 Sequence 4, Appll
981 23 60.5 3567 3 US-08-439-009A-4 Sequence 4, Appll
982 23 60.5 3729 2 US-08-804-227C-4 Sequence 4, Appll
983 23 60.5 4551 3 US-09-320-878-1 Sequence 1, Appll
984 23 60.5 4613 4 US-09-105-537-31 Sequence 31, Appll
985 23 60.5 7257 3 US-09-335-409-5 Sequence 5, Appll
986 23 60.5 7257 4 US-09-568-102-5 Sequence 5, Appll
987 23 60.5 7257 4 US-09-567-969-5 Sequence 5, Appll
988 23 60.5 7257 4 US-09-568-480-5 Sequence 5, Appll
989 23 60.5 7257 4 US-09-568-480-5 Sequence 5, Appll
990 23 60.5 7257 4 US-09-568-472-5 Sequence 5, Appll
991 23 60.5 7257 4 US-09-567-899-5 Sequence 5, Appll
992 23 60.5 11877 4 US-09-105-537-6 Sequence 6, Appll
993 22 57.9 7 3 US-08-556-419-14 Sequence 14, Appll
994 22 57.9 11 2 US-08-726-136-13 Sequence 13, Appll
995 22 57.9 11 4 US-09-103-434-13 Sequence 13, Appll
996 22 57.9 14 1 US-09-687-594-13 Sequence 13, Appll
997 22 57.9 14 1 US-08-366-276-9 Sequence 9, Appll
998 22 57.9 18 3 US-08-940-095-237 Sequence 237, App
999 22 57.9 18 3 US-08-940-093-237 Sequence 237, App
1000 22 57.9 18 3 US-08-940-096-237 Sequence 237, App

```

## ALIGNMENTS

```

RESULT 1
US-09-428-517-2
; Sequence 2, Application US/09428517
; Patent No. 6251636
; GENERAL INFORMATION:
; APPLICANT: Shah, Sanjay Krishna Kant
; APPLICANT: Medanlel, Robert
; APPLICANT: Tang, Li
; TITLE OF INVENTION: RECOMBINANT OLEANDOLIDE POLYKETIDE SYNTHASE
; FILE REFERENCE: 30062-20029.00
; CURRENT APPLICATION NUMBER: US/09/428,517
; CURRENT FILING DATE: 1999-10-28
; EARLIER APPLICATION NUMBER: 60/120,254
; EARLIER FILING DATE: 1999-02-16
; EARLIER APPLICATION NUMBER: 60/106,100
; EARLIER FILING DATE: 1998-10-29
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4150
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Recombinant
US-09-428-517-2
; OTHER INFORMATION: Oleandolide PKS

Query Match      84.2%; Score 32; DB 4; Length 4150;
Best Local Similarity 87.5%; Pred. No. 4,1e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 LKDDLEA 9
DB 1781 LRDDLEA 1788

RESULT 2
US-09-335-409-3
; Sequence 3, Application US/09335409
; Patent No. 6121029
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James

```

```

; APPLICANT: Molnar, Istvan
; APPLICANT: Zirkle, Ross
; APPLICANT: Cyr, Devon
; APPLICANT: Goerlach, Joern
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILLONES
; FILE REFERENCE: 4-30582A
; CURRENT APPLICATION NUMBER: US/09/335,409
; CURRENT FILING DATE: 1999-06-17
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1410
; TYPE: PRT
; ORGANISM: Sorangium cellulosum
US-09-335-409-3

Query Match      78.9%; Score 30; DB 3; Length 1410;
Best Local Similarity 75.0%; Pred. No. 3,2e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LKDDLEA 9
DB 474 LDDMLEA 481

```

```

RESULT 3
US-09-568-102-3
; Sequence 3, Application US/09568102
; Patent No. 6346404
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James
; APPLICANT: Molnar, Istvan
; APPLICANT: Zirkle, Ross
; APPLICANT: Cyr, Devon
; APPLICANT: Goerlach, Joern
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILLONES
; FILE REFERENCE: 4-30582A
; CURRENT APPLICATION NUMBER: US/09/568,102
; CURRENT FILING DATE: 2000-05-10
; PRIOR APPLICATION NUMBER: 09/335,409
; PRIOR FILING DATE: 1999-06-17
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1410
; TYPE: PRT
; ORGANISM: Sorangium cellulosum
US-09-568-102-3

Query Match      78.9%; Score 30; DB 4; Length 1410;
Best Local Similarity 75.0%; Pred. No. 3,2e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LKDDLEA 9
DB 474 LDDMLEA 481

RESULT 4
US-09-567-969-3
; Sequence 3, Application US/09567969
; Patent No. 6355457
; GENERAL INFORMATION:
; APPLICANT: Schupp, Thomas
; APPLICANT: Ligon, James
; APPLICANT: Molnar, Istvan
; APPLICANT: Zirkle, Ross
; APPLICANT: Cyr, Devon
; APPLICANT: Goerlach, Joern
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILLONES
; FILE REFERENCE: 4-30582A
; CURRENT APPLICATION NUMBER: US/09/567,969

```



;; CURRENT FILING DATE: 2000-05-10  
;; PRIOR APPLICATION NUMBER: 09/335,409  
;; PRIOR FILING DATE: 1999-06-17  
;; NUMBER OF SEQ ID NOS: 30  
;; SOFTWARE: Patentln Ver. 2.0  
;; SEQ ID NO 3  
;; LENGTH: 1410  
;; TYPE: PRT  
;; ORGANISM: Sorangium cellulosum  
US-09-567-969-3

Query Match 78.9%; Score 30; DB 4; Length 1410;  
Best Local Similarity 75.0%; Pred. No. 3.2e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LXDLLA 9  
1 11:111  
Db 474 LDDMLA 481

RESULT 5  
US-09-568-480-3  
; Sequence 3, Application US/09568480  
; Patent No. 6355458  
; GENERAL INFORMATION:  
; APPLICANT: Schupp, Thomas  
; APPLICANT: Ligon, James  
; APPLICANT: Molnar, Istvan  
; APPLICANT: Zirkle, Ross  
; APPLICANT: Cyr, Devon  
; APPLICANT: Goelach, Joern  
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES  
; FILE REFERENCE: 4-30582A  
; CURRENT APPLICATION NUMBER: US/09/568,480  
; CURRENT FILING DATE: 2000-05-10  
; PRIOR APPLICATION NUMBER: 09/335,409  
; PRIOR FILING DATE: 1999-06-17  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: Patentln Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 1410  
; TYPE: PRT  
; ORGANISM: Sorangium cellulosum  
US-09-568-480-3

Query Match 78.9%; Score 30; DB 4; Length 1410;  
Best Local Similarity 75.0%; Pred. No. 3.2e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LXDLLA 9  
1 11:111  
Db 474 LDDMLA 481

RESULT 6  
US-09-568-486-3  
; Sequence 3, Application US/09568486  
; Patent No. 6355459  
; GENERAL INFORMATION:  
; APPLICANT: Schupp, Thomas  
; APPLICANT: Ligon, James  
; APPLICANT: Molnar, Istvan  
; APPLICANT: Zirkle, Ross  
; APPLICANT: Cyr, Devon  
; APPLICANT: Goelach, Joern  
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES  
; FILE REFERENCE: 4-30582A  
; CURRENT APPLICATION NUMBER: US/09/568,486  
; CURRENT FILING DATE: 2000-05-10  
; PRIOR APPLICATION NUMBER: 09/335,409  
; PRIOR FILING DATE: 1999-06-17  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: Patentln Ver. 2.0

;; SEQ ID NO 3  
;; LENGTH: 1410  
;; TYPE: PRT  
;; ORGANISM: Sorangium cellulosum  
US-09-568-486-3

Query Match 78.9%; Score 30; DB 4; Length 1410;  
Best Local Similarity 75.0%; Pred. No. 3.2e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LXDLLA 9  
1 11:111  
Db 474 LDDMLA 481

RESULT 7  
US-09-568-472-3  
; Sequence 3, Application US/09568472  
; Patent No. 6358719  
; GENERAL INFORMATION:  
; APPLICANT: Schupp, Thomas  
; APPLICANT: Ligon, James  
; APPLICANT: Molnar, Istvan  
; APPLICANT: Zirkle, Ross  
; APPLICANT: Cyr, Devon  
; APPLICANT: Goelach, Joern  
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES  
; FILE REFERENCE: 4-30582A  
; CURRENT APPLICATION NUMBER: US/09/568,472  
; CURRENT FILING DATE: 2000-05-10  
; PRIOR APPLICATION NUMBER: 09/335,409  
; PRIOR FILING DATE: 1999-06-17  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: Patentln Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 1410  
; TYPE: PRT  
; ORGANISM: Sorangium cellulosum  
US-09-568-472-3

Query Match 78.9%; Score 30; DB 4; Length 1410;  
Best Local Similarity 75.0%; Pred. No. 3.2e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LXDLLA 9  
1 11:111  
Db 474 LDDMLA 481

RESULT 8  
US-09-567-899-3  
; Sequence 3, Application US/09567899  
; Patent No. 6383787  
; GENERAL INFORMATION:  
; APPLICANT: Schupp, Thomas  
; APPLICANT: Ligon, James  
; APPLICANT: Molnar, Istvan  
; APPLICANT: Zirkle, Ross  
; APPLICANT: Cyr, Devon  
; APPLICANT: Goelach, Joern  
; TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES  
; FILE REFERENCE: 4-30582A  
; CURRENT APPLICATION NUMBER: US/09/567,899  
; CURRENT FILING DATE: 2000-05-10  
; PRIOR APPLICATION NUMBER: 09/335,409  
; PRIOR FILING DATE: 1999-06-17  
; NUMBER OF SEQ ID NOS: 30  
; SOFTWARE: Patentln Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 1410  
; TYPE: PRT  
; ORGANISM: Sorangium cellulosum  
US-09-567-899-3

Query Match 78.9%; Score 30; DB 4; Length 1410;  
 Best Local Similarity 75.0%; Pred. No. 3.2e+02;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LKDDLEA 9  
 111111  
 Db 474 LKDDLEA 481

RESULT 9  
 PCT-US95-12357A-1  
 ; Sequence 1, Application PC/TUS9512357A  
 ; GENERAL INFORMATION:  
 ; APPLICANT: St. Jude Children's Research Hospital  
 ; TITLE OF INVENTION: Nucleic Acid Encoding Mutant Matrix  
 ; TITLE OF INVENTION: Proteins Useful for Attenuation of Enhancement of  
 ; TITLE OF INVENTION: Influenza A Virus, Vaccines and Methods of making and  
 ; NUMBER OF SEQUENCES: 13  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX  
 ; STREET: 1100 New York Avenue, N.W.  
 ; CITY: Washington  
 ; STATE: DC  
 ; COUNTRY: USA  
 ; ZIP: 20005  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: PCT/US95/12357A  
 ; FILING DATE: 29-SEPT-1995  
 ; CLASSIFICATION:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/316,419  
 ; FILING DATE: 30-SEPT-1994  
 ; CLASSIFICATION:  
 ; APPLICATION NUMBER: 08/471,100  
 ; FILING DATE: 6-JUNE-1995  
 ; CLASSIFICATION:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Fox, Samuel L.  
 ; REGISTRATION NUMBER: 30,353  
 ; REFERENCE/DOCKET NUMBER: 0656.048PC01  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 202-371-2600  
 ; TELEFAX: 202-371-2540  
 ; INFORMATION FOR SEQ ID NO: 1:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 251 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: not relevant  
 ; TOPOLOGY: not relevant  
 ; MOLECULE TYPE: peptide  
 ; PCT-US95-12357A-1

Query Match 73.7%; Score 28; DB 5; Length 251;  
 Best Local Similarity 85.7%; Pred. No. 1.3e+02;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 LKDDLEA 8  
 111111  
 Db 228 LKDDLEA 234

RESULT 10  
 US-08-674-168-29  
 ; Sequence 29, Application US/08674168  
 ; Patent No. 5804414  
 ; GENERAL INFORMATION:

APPLICANT: MORIYA, Mika  
 APPLICANT: MATSUI, Hiroshi  
 APPLICANT: YOKOZAKI, Kenzo  
 APPLICANT: HIRANO, Seiko  
 APPLICANT: HAYAKAWA, Atsushi  
 APPLICANT: IZUT, Masako  
 APPLICANT: SOGIMOTO, Masakazu  
 TITLE OF INVENTION: METHOD OF AMPLIFYING GENE USING  
 TITLE OF INVENTION: ARTIFICIAL TRANSPOSON  
 NUMBER OF SEQUENCES: 32  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
 P.C.  
 STREET: 1755 JEFFERSON DAVIS HIGHWAY, SUITE # 400  
 CITY: ARLINGTON  
 STATE: VIRGINIA  
 COUNTRY: USA  
 ZIP: 22202  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/674,168  
 ; FILING DATE: 01-JUL-1996  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: JP 7-166541  
 ; FILING DATE: 30-JUN-1995  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: OBLON, NORMAN F.  
 ; REGISTRATION NUMBER: 24,618  
 ; REFERENCE/DOCKET NUMBER: 10-810-0  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (703) 413-3000  
 ; TELEFAX: (703) 413-2220  
 ; TELEX: 248855 OPAT UR  
 ; INFORMATION FOR SEQ ID NO: 29:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 550 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-08-674-168-29

Query Match 73.7%; Score 28; DB 1; Length 550;  
 Best Local Similarity 66.7%; Pred. No. 3e+02;  
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 VLKDDLEA 9  
 111111  
 Db 383 VLKDDLEA 391

RESULT 11  
 US-08-985-908-19  
 ; Sequence 19, Application US/08985908  
 ; Patent No. 6004773  
 ; GENERAL INFORMATION:  
 ; APPLICANT: MASAYUKI ARAKI, MASAKAZU SUGIMOTO, YASUHIKO YOSHIMURA, AND TSUYOSHI  
 ; TITLE OF INVENTION: METHOD FOR PRODUCING L-LYSINE  
 ; NUMBER OF SEQUENCES: 31  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT, P.C.  
 ; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR  
 ; CITY: ARLINGTON  
 ; STATE: VIRGINIA  
 ; COUNTRY: USA  
 ; ZIP: 22202  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible

```

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/985,908
FILING DATE: 05-DEC-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 8-325659
FILING DATE: 05-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: NORMAN F. OBLON
REGISTRATION NUMBER: 24,618
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 550 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-985-908-19

Query Match
Best Local Similarity 73.7%; Score 28; DB 3; Length 550;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 VLXDDLLEA 9
Db 383 VTLDLVEA 391

RESULT 12
US-08-852-730-4
Sequence 4, Application US/08852730
Patent No. 6090597
GENERAL INFORMATION:
APPLICANT: SEIKO HIRANO, MASAKAZU SUGIMOTO, EIICHI NAKANO,
APPLICANT: MASAKO IZUI, ATSUSHI HAYAKAWA, YASUHIKO YOSHIHARA, AND TSUYOSHI
APPLICANT: NAKAMATSU
TITLE OF INVENTION: METHOD OF PRODUCING L-LYSINE
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER AND NEUSTADT, P.C.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY
CITY: ARLINGTON
STATE: VA
ZIP: 22026
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/852,730
FILING DATE: 05-07-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 8-142812
FILING DATE: 05-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: NORMAN F. OBLON
REGISTRATION NUMBER: 24,618
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 550 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-852-730-4

```

```

Query Match
Best Local Similarity 73.7%; Score 28; DB 3; Length 550;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 VLXDDLLEA 9
Db 383 VTLDLVEA 391

RESULT 13
US-08-985-916-11
Sequence 11, Application US/08985916
Patent No. 6221636
GENERAL INFORMATION:
APPLICANT: ATSUSHI HAYAKAWA, MASAKAZU SUGIMOTO, YASUHIKO YOSHIHARA, AND TSUYOSH
TITLE OF INVENTION: METHOD FOR PRODUCING L-LYSINE
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT, P.C.
STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, FOURTH FLOOR
CITY: ARLINGTON
STATE: VA
ZIP: 22152
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/985,916
FILING DATE: 05-DEC-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 8-325658
FILING DATE: 05-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: NORMAN F. OBLON
REGISTRATION NUMBER: 24,618
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 550 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-985-916-11

Query Match
Best Local Similarity 73.7%; Score 28; DB 4; Length 550;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 VLXDDLLEA 9
Db 383 VTLDLVEA 391

RESULT 14
US-08-416-603-4
Sequence 4, Application US/08416603
Patent No. 5866780
GENERAL INFORMATION:
APPLICANT: Law, Marcus
APPLICANT: Hepburn, Ledare
APPLICANT: Reddick, Bradford B.
TITLE OF INVENTION: Maize Chlorotic Dwarf Virus Genome and
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Saliwanchik & Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1

```

CITY: Gainesville  
STATE: FL  
COUNTRY: USA  
ZIP: 32606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.30B  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/416,603  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Lloyd, Jeffrey  
REGISTRATION NUMBER: 35,589  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 904-375-8100  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3457 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-416-603-4

Query Match 73.7%; Score 28; DB 2; Length 3457;  
Best Local Similarity 55.6%; Pred. No. 2.3e+03;  
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LXXDLEA 9  
Db 1772 MYVDLDA 1780

RESULT 15  
US-09-428-517-4  
Sequence 4, Application US/09428517  
Patent No. 6251636  
GENERAL INFORMATION:  
APPLICANT: Bellach, Mary C.  
APPLICANT: Shah, Sanjay Krishnakant  
APPLICANT: McDaniel, Robert  
APPLICANT: Tang, Li  
TITLE OF INVENTION: RECOMBINANT OLEANDOLIDE POLYKETIDE SYNTHASE  
FILE REFERENCE: 30062-20029.00  
CURRENT FILING DATE: 1999-10-28  
EARLIER FILING DATE: 1999-02-16  
EARLIER APPLICATION NUMBER: 60/120,254  
EARLIER FILING DATE: 1999-02-16  
EARLIER APPLICATION NUMBER: 60/106,100  
NUMBER OF SEQ ID NOS: 12  
SOFTWARE: Patentln Ver. 2.1  
SEQ ID NO 4  
LENGTH: 3519  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Recombinant  
OTHER INFORMATION: Oleandolide PKS  
US-09-428-517-4

Query Match 73.7%; Score 28; DB 4; Length 3519;  
Best Local Similarity 75.0%; Pred. No. 2.4e+03;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LXXDLEA 9  
Db 2421 LRDELLEA 2428

RESULT 16

US-09-428-517-3  
Sequence 3, Application US/09428517  
Patent No. 6251636  
GENERAL INFORMATION:  
APPLICANT: Bellach, Mary C.  
APPLICANT: Shah, Sanjay Krishnakant  
APPLICANT: McDaniel, Robert  
APPLICANT: Tang, Li  
TITLE OF INVENTION: RECOMBINANT OLEANDOLIDE POLYKETIDE SYNTHASE  
FILE REFERENCE: 30062-20029.00  
CURRENT FILING DATE: 1999-10-28  
EARLIER FILING DATE: 1999-02-16  
EARLIER APPLICATION NUMBER: 60/120,254  
EARLIER FILING DATE: 1999-02-16  
EARLIER APPLICATION NUMBER: 60/106,100  
NUMBER OF SEQ ID NOS: 12  
SOFTWARE: Patentln Ver. 2.1  
SEQ ID NO 3  
LENGTH: 3816  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Recombinant  
OTHER INFORMATION: Oleandolide PKS  
US-09-428-517-3

Query Match 73.7%; Score 28; DB 4; Length 3816;  
Best Local Similarity 75.0%; Pred. No. 2.6e+03;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LXXDLEA 9  
Db 2327 LRDELLEA 2334

RESULT 17  
US-08-899-330-6  
Sequence 6, Application US/08899330  
Patent No. 6177275  
GENERAL INFORMATION:  
APPLICANT: CORUZZI, GLORIA  
APPLICANT: LAM, HON-MING  
APPLICANT: HSIEH, MING-HSIUN  
TITLE OF INVENTION: PLANT NITROGEN REGULATORY  
NUMBER OF SEQUENCES: 21  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: PENNIE & EDMONDS LLP  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/899,330  
FILING DATE: 23-JUL-1997  
CLASSIFICATION: 800  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/022,328  
FILING DATE: 24-JUL-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 5914-042-999  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212)7909090  
TELEFAX: (212)8699741

TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 111 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-899-330-6

Query Match 71.1%; Score 27; DB 4; Length 111;  
Best Local Similarity 62.3%; Pred. No. 82;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 VLXDDL 8  
: | | | | |  
Db 63 VIGDDLV 70

RESULT 18  
US-08-480-640A-192  
; Sequence 192, Application US/08480640A  
; Patent No. 6033904  
; GENERAL INFORMATION:  
; APPLICANT: Cochran, Mark D.  
; APPLICANT: Junker, David E.  
; TITLE OF INVENTION: Recombinant Swinepox Virus  
; NUMBER OF SEQUENCES: 225  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: John P. White  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/480,640A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P.  
REGISTRATION NUMBER: 28,678  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 278-0400  
TELEFAX: (212) 391-0525  
INFORMATION FOR SEQ ID NO: 192:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 313 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-480-640A-192

Query Match 71.1%; Score 27; DB 3; Length 313;  
Best Local Similarity 71.4%; Pred. No. 2.6e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 VLXDDL 7  
: | | | | |  
Db 184 ILSDLL 190

RESULT 19  
US-08-686-968C-192  
; Sequence 192, Application US/08686968C  
; Patent No. 6221361  
; GENERAL INFORMATION:  
; APPLICANT: Cochran, Mark D.

APPLICANT: Junker, David E.  
TITLE OF INVENTION: Recombinant Swinepox Virus  
FILE REFERENCE: 39119-H/JML  
CURRENT APPLICATION NUMBER: US/08/686,968C  
CURRENT FILING DATE: 1996-07-25  
NUMBER OF SEQ ID NOS: 231  
SOFTWARE: Patentln Ver. 2.1  
SEQ ID NO 192  
LENGTH: 313  
TYPE: PR1  
ORGANISM: Swinepox virus  
US-08-686-968C-192

Query Match 71.1%; Score 27; DB 4; Length 313;  
Best Local Similarity 71.4%; Pred. No. 2.6e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 VLXDDL 7  
: | | | | |  
Db 184 ILSDLL 190

RESULT 20  
US-08-488-237A-192  
; Sequence 192, Application US/08488237A  
; Patent No. 6251403  
; GENERAL INFORMATION:  
; APPLICANT: Cochran, Mark D.  
; APPLICANT: Junker, David E.  
; TITLE OF INVENTION: Recombinant Swinepox Virus  
; NUMBER OF SEQUENCES: 225  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: John P. White  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/488,237A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P.  
REGISTRATION NUMBER: 28,678  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 278-0400  
TELEFAX: (212) 391-0525  
INFORMATION FOR SEQ ID NO: 192:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 313 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-488-237A-192

Query Match 71.1%; Score 27; DB 4; Length 313;  
Best Local Similarity 71.4%; Pred. No. 2.6e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 VLXDDL 7  
: | | | | |  
Db 184 ILSDLL 190

RESULT 21  
US-08-375-992A-192  
; Sequence 192, Application US/08375992A

Patent No. 6328975  
GENERAL INFORMATION:  
APPLICANT: Cochran, Mark D.  
APPLICANT: Junker, David E.  
TITLE OF INVENTION: Recombinant Swinepox Virus  
NUMBER OF SEQUENCES: 220  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: John P. White  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/375,992A  
FILING DATE: Herewith  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P  
REGISTRATION NUMBER: 28,678  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 278-0400  
TELEFAX: (212) 391-0525  
INFORMATION FOR SEQ ID NO: 192:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 313 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-375-992A-192

Query Match 71.1%; Score 27; DB 4; Length 313;  
Best Local Similarity 71.4%; Pred. No. 2.6e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLXDDL 7  
DB 184 ILSDDL 190

RESULT 22  
US-09-154-750A-89  
Sequence 89, Application US/09154750A  
Patent No. 6432640  
GENERAL INFORMATION:  
APPLICANT: Vogelstein, Bert  
APPLICANT: Kinzler, Kenneth  
APPLICANT: Polyak, Kornelia  
TITLE OF INVENTION: p53-Induced Apoptosis  
FILE REFERENCE: 1107.75357  
CURRENT APPLICATION NUMBER: US/09/154,750A  
CURRENT FILING DATE: 1998-09-17  
PRIOR APPLICATION NUMBER: 60/059,153  
PRIOR FILING DATE: 1997-09-17  
PRIOR APPLICATION NUMBER: 60/079817  
PRIOR FILING DATE: 1998-03-30  
NUMBER OF SEQ ID NOS: 93  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 89  
LENGTH: 324  
TYPE: PRT  
ORGANISM: Vigna  
US-09-154-750A-89

Query Match 71.1%; Score 27; DB 4; Length 324;  
Best Local Similarity 66.7%; Pred. No. 2.7e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 VLXDDLEA 9  
DB 270 VTRDELLEA 278

RESULT 23  
US-09-222-938A-31  
Sequence 31, Application US/09222938A  
Patent No. 6437108  
GENERAL INFORMATION:  
APPLICANT: Youngman, Phillip  
APPLICANT: Fritz, Christian  
APPLICANT: Murphy, Christopher  
APPLICANT: Guzman, Luz-Maria  
TITLE OF INVENTION: ESSENTIAL BACTERIAL GENES AND THEIR USE  
FILE REFERENCE: 07334/060001  
CURRENT APPLICATION NUMBER: US/09/222,938A  
CURRENT FILING DATE: 1998-12-30  
NUMBER OF SEQ ID NOS: 102  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 31  
LENGTH: 324  
TYPE: PRT  
ORGANISM: Streptococcus pneumoniae  
US-09-222-938A-31

Query Match 71.1%; Score 27; DB 4; Length 324;  
Best Local Similarity 71.4%; Pred. No. 2.7e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LXDLE 8  
DB 64 TRDDLE 70

RESULT 24  
US-09-562-737-7  
Sequence 7, Application US/09562737  
Patent No. 6428967  
GENERAL INFORMATION:  
APPLICANT: Hertz, Joachim  
APPLICANT: Gotthardt, Michael  
TITLE OF INVENTION: LDL Receptor Signaling Pathways  
FILE REFERENCE: UTSW0708  
CURRENT APPLICATION NUMBER: US/09/562,737  
CURRENT FILING DATE: 2000-05-01  
NUMBER OF SEQ ID NOS: 132  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 7  
LENGTH: 333  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Synthetic  
US-09-562-737-7

Query Match 71.1%; Score 27; DB 4; Length 333;  
Best Local Similarity 62.5%; Pred. No. 2.8e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 LXDLEA 9  
DB 267 MVDDLEA 274

RESULT 25  
US-08-480-640A-114  
Sequence 114, Application US/08480640A  
Patent No. 6033904  
GENERAL INFORMATION:  
APPLICANT: Cochran, Mark D.  
APPLICANT: Junker, David E.

TITLE OF INVENTION: Recombinant Swinepox Virus  
NUMBER OF SEQUENCES: 225  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: John P. White  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/480,640A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P  
REGISTRATION NUMBER: 28,678  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 278-0400  
TELEFAX: (212) 391-0525  
INFORMATION FOR SEQ ID NO: 114:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 389 amino acids  
TYPE: amino acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Swinepox virus  
STRAIN: Kasza  
INDIVIDUAL ISOLATE: S-SPV-001  
IMMEDIATE SOURCE:  
CLONE: 515-85.1  
POSITION IN GENOME:  
MAP POSITION: ~23.2  
UNITS: %G  
US-08-480-640A-114

Query Match 71.1%; Score 27; DB 3; Length 389;  
Best Local Similarity 71.4%; Pred. No. 3.3e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLXDDL 7  
: | | | |  
Db 260 ILSDLL 266

RESULT 26  
US-08-295-802-114  
; Sequence 114, Application US/08295802  
; Patent No. 6127163  
; GENERAL INFORMATION:  
; APPLICANT: Cochran Ph.D., Mark D  
; APPLICANT: Junker M.S., David E  
; TITLE OF INVENTION: Recombinant Swinepox Virus  
; NUMBER OF SEQUENCES: 188  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: John P. White  
; STREET: 30 Rockefeller Plaza  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10112  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentln Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/295,802  
FILING DATE: Herewith  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212)977-9550  
TELEFAX: (212)664-0525  
TELEX: 422523  
INFORMATION FOR SEQ ID NO: 114:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 389 amino acids  
TYPE: amino acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Swinepox virus  
STRAIN: Kasza  
INDIVIDUAL ISOLATE: S-SPV-001  
IMMEDIATE SOURCE:  
CLONE: 515-85.1  
POSITION IN GENOME:  
MAP POSITION: ~23.2  
UNITS: %G  
US-08-295-802-114

Query Match 71.1%; Score 27; DB 3; Length 389;  
Best Local Similarity 71.4%; Pred. No. 3.3e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLXDDL 7  
: | | | |  
Db 260 ILSDLL 266

RESULT 27  
US-08-488-237A-114  
; Sequence 114, Application US/08488237A  
; Patent No. 6251403  
; GENERAL INFORMATION:  
; APPLICANT: Cochran, Mark D.  
; APPLICANT: Junker, David E.  
; TITLE OF INVENTION: Recombinant Swinepox Virus  
; NUMBER OF SEQUENCES: 225  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: John P. White  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentln Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/488,237A  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: White, John P  
; REGISTRATION NUMBER: 28,678  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 278-0400  
; TELEFAX: (212) 391-0525  
; INFORMATION FOR SEQ ID NO: 114:  
; SEQUENCE CHARACTERISTICS:

LENGTH: 389 amino acids  
TYPE: amino acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Swinepox virus  
STRAIN: Kasza  
INDIVIDUAL ISOLATE: S-SPV-001  
IMMEDIATE SOURCE:  
CLONE: 515-85.1  
POSITION IN GENOME:  
MAP POSITION: -23.2  
UNITS: %G  
US-08-488-237A-114

Query Match 71.1%; Score 27; DB 4; Length 389;  
Best Local Similarity 71.4%; Pred. No. 3.3e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLXDDL 7  
: 1 1 1 1  
Db 260 ILSDLL 266

RESULT 28  
US-08-375-992A-114  
Sequence 114, Application US/08375992A  
Patent No. 6328975  
GENERAL INFORMATION:  
APPLICANT: Cochran, Mark D.  
TITLE OF INVENTION: Recombinant Swinepox Virus  
NUMBER OF SEQUENCES: 220  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: John P. White  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/375,992A  
FILING DATE: Herewith  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P  
REGISTRATION NUMBER: 28,678  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 278-0400  
TELEFAX: (212) 391-0525  
INFORMATION FOR SEQ ID NO: 114:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 389 amino acids  
TYPE: amino acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Swinepox virus  
STRAIN: Kasza  
INDIVIDUAL ISOLATE: S-SPV-001  
IMMEDIATE SOURCE:  
CLONE: 515-85.1

POSITION IN GENOME:  
MAP POSITION: ~23.2  
UNITS: %G  
US-08-375-992A-114

Query Match 71.1%; Score 27; DB 4; Length 389;  
Best Local Similarity 71.4%; Pred. No. 3.3e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLXDDL 7  
: 1 1 1 1  
Db 260 ILSDLL 266

RESULT 29  
US-09-167-299-5  
Sequence 5, Application US/09167299  
Patent No. 624539  
GENERAL INFORMATION:  
APPLICANT: Kron, Michael  
APPLICANT: Hartlein, Michael  
APPLICANT: Michito, Harikata  
TITLE OF INVENTION: Human Asparaginyl-tRNA Synthetase DNA  
FILE REFERENCE: Aparaginy1-tRNA Synthetase  
CURRENT APPLICATION NUMBER: US/09/167,299  
CURRENT FILING DATE: 1998-10-06  
NUMBER OF SEQ ID NOS: 7  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 5  
LENGTH: 438  
TYPE: PRT  
ORGANISM: Thermus aquaticus  
PUBLICATION INFORMATION:  
DATABASE ACCESSION NUMBER: X91009  
DATABASE ENTRY DATE: 1996-08-21  
US-09-167-299-5

Query Match 71.1%; Score 27; DB 4; Length 438;  
Best Local Similarity 85.7%; Pred. No. 3.8e+02;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLXDDL 7  
: 1 1 1 1  
Db 348 VLNDLL 354

RESULT 30  
US-09-409-180A-1  
Sequence 1, Application US/09409180A  
Patent No. 6444802  
GENERAL INFORMATION:  
APPLICANT: Kapeller-Libermann, Rosana  
APPLICANT: White, David  
TITLE OF INVENTION: 22196, A No. 6444802el Human Amino-peptidase  
FILE REFERENCE: 5800-59  
CURRENT APPLICATION NUMBER: US/09/409,180A  
CURRENT FILING DATE: 1999-09-30  
NUMBER OF SEQ ID NOS: 3  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 1  
LENGTH: 704  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-409-180A-1

Query Match 71.1%; Score 27; DB 4; Length 704;  
Best Local Similarity 71.4%; Pred. No. 6.4e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LXDLL 8  
: 1 1 1 1  
Db 555 IADLL 561



RESULT 31  
PCT-US92-00731-13  
Sequence 13, Application PC/TUS9200731  
GENERAL INFORMATION:  
APPLICANT: Kawabata, Shunichiro  
APPLICANT: Davie, Earl W.  
TITLE OF INVENTION: MICROSOFT ENDOPEPTIDASE  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESS: Seed and Berry  
STREET: 6300 Columbia Center  
City: Seattle  
STATE: Washington  
COUNTRY: US  
ZIP: 98104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US92/00731  
FILING DATE: 19920128  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/646,997  
FILING DATE: 28-JAN-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: MAKI, David J.  
REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 990008.415PC  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206)622-4900  
TELEFAX: (206)682-6031  
TELEX: 3723836  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 709 amino acids  
TYPE: AMINO ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: N-terminal  
PCT-US92-00731-13

Query Match 71.1%; Score 27; DB 5; Length 709;  
Best Local Similarity 71.4%; Pred. No. 6.5e+02;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 LKDDLE 8  
DB 609 IADDLE 615

RESULT 32  
US-07-935-311A-4  
Sequence 4, Application US/07935311A  
Patent No. 5378609  
GENERAL INFORMATION:  
APPLICANT: Di Fiore, Pier Paolo  
APPLICANT: Fazioli, Francesca  
TITLE OF INVENTION: Substrate of the Epidermal Growth  
TITLE OF INVENTION: Factor Kinase  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Knobbe, Martens, Olson & Bear  
STREET: 620 Newport Center Drive, 16th Floor  
CITY: Newport Beach  
STATE: CA  
COUNTRY: USA  
ZIP: 92660

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/935,311A  
FILING DATE: 19920825  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Israelson, Ned A.  
REGISTRATION NUMBER: 29,655  
REFERENCE/DOCKET NUMBER: NIH035.001A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 235-0176  
TELEFAX: (619) 235-0176  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 821 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-935-311A-4

Query Match 71.1%; Score 27; DB 1; Length 821;  
Best Local Similarity 62.5%; Pred. No. 7.6e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VKDDLE 8  
DB 551 VKDDLE 558

RESULT 33  
US-08-368-079-4  
Sequence 4, Application US/08368079  
Patent No. 5610018  
GENERAL INFORMATION:  
APPLICANT: Di Fiore, Pier Paolo  
APPLICANT: Fazioli, Francesca  
TITLE OF INVENTION: eps8, A Substrate for the Epidermal Growth Factor  
TITLE OF INVENTION: Kinase, Antibodies Thereof, and Methods of Use Thereof  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Knobbe, Martens, Olson & Bear  
STREET: 620 Newport Center Drive, 16th Floor  
CITY: Newport Beach  
STATE: CA  
COUNTRY: USA  
ZIP: 92660  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/368,079  
FILING DATE: 03-JAN-1995  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/935,311  
FILING DATE: 25-AUG-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Israelson, Ned A.  
REGISTRATION NUMBER: 29,655  
REFERENCE/DOCKET NUMBER: NIH035.001DV1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 235-8550  
TELEFAX: (619) 235-0176  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 821 amino acids  
TYPE: amino acid

TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-368-079-4

Query Match 71.1%; Score 27; DB 1; Length 821;  
Best Local Similarity 62.5%; Pred. No. 7.6e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 VLXDDLE 8  
1: 11:11  
DB 551 VKDDVLE 558

RESULT 34  
PCT-US93-07996-4  
Sequence 4, Application PC/TUS9307996  
GENERAL INFORMATION:  
APPLICANT: The Government of the United States, as represented by the  
APPLICANT: Secretary of Health and Human Services  
TITLE OF INVENTION: Substrate of the Epidermal Growth Factor Kinase  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Knodde, Martens, Olson & Bear  
STREET: 620 Newport Center Drive, 16th Floor  
CITY: Newport Beach  
STATE: CA  
COUNTRY: USA  
ZIP: 92660

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US93/07996  
FILING DATE: 19930825  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 235-8550  
TELEFAX: (619) 235-0176  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 821 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US93-07996-4

Query Match 71.1%; Score 27; DB 5; Length 821;  
Best Local Similarity 62.5%; Pred. No. 7.6e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 VLXDDLE 8  
1: 11:11  
DB 551 VKDDVLE 558

RESULT 35  
US-08-680-326-41  
Sequence 41, Application US/08680326  
Patent No. 5925733  
GENERAL INFORMATION:  
APPLICANT: ROSE, TIMOTHY M.  
APPLICANT: BOSCH, MARINIX  
APPLICANT: STRAND, KURT  
APPLICANT: TODARO, GEORGE J.  
TITLE OF INVENTION: DNA POLYMERASE OF GAMMA HERPES VIRUSES  
TITLE OF INVENTION: ASSOCIATED WITH KAPOSI'S SARCOMA AND RETROPERITONEAL  
NUMBER OF SEQUENCES: 152  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORRISON & FOERSTER  
STREET: 755 Page Mill Road  
CITY: Palo Alto

STATE: California  
COUNTRY: USA  
ZIP: 94304-1018

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/680,326  
FILING DATE:  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Schiff, J. Michael  
REGISTRATION NUMBER: 40,253  
REFERENCE/DOCKET NUMBER: 29938-20001.00  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 813-5600  
TELEFAX: (415) 494-0792  
TELEX: 706141

INFORMATION FOR SEQ ID NO: 41:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 985 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-680-326-41

Query Match 71.1%; Score 27; DB 2; Length 985;  
Best Local Similarity 62.5%; Pred. No. 9.3e+02;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 VLXDDLE 8  
1: 11:11  
DB 866 VLDDVLE 873

RESULT 36  
US-08-452-083-2  
Sequence 2, Application US/08452083  
Patent No. 5756327  
GENERAL INFORMATION:  
APPLICANT: Sassanfar, Mandana  
APPLICANT: Schimmel, Paul R.  
TITLE OF INVENTION: RECOMBINANT MYCOBACTERIAL ISOLEUCYL-tRNA  
NUMBER OF SEQUENCES: 27  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
STREET: Two Millitia Drive  
CITY: Lexington  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02173  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/452,083  
FILING DATE: 26-MAY-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/305,765  
FILING DATE: 13-SEP-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Brook, David E.  
REGISTRATION NUMBER: 22,592  
REFERENCE/DOCKET NUMBER: CPI94-088  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 861-6240  
TELEFAX: (617) 861-9540

;; INFORMATION FOR SEQ ID NO: 2:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 1045 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
US-08-452-083-2

Query Match 71.1%; Score 27; DB 1; Length 1045;  
Best Local Similarity 66.7%; Pred. No. 9.9e+02;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 VLXDDLEA 9  
111111  
Db 720 LVDDLES 728

RESULT 37  
US-08-851-843A-141  
; Sequence 141, Application US/08851843A  
; Patent No. 6093809  
; GENERAL INFORMATION:  
; APPLICANT: Cech, Thomas R.  
; APPLICANT: Lingner, Joachim  
; APPLICANT: Nakamura, Toru  
; APPLICANT: Chapman, Karen B.  
; APPLICANT: Morin, Gregg B.  
; APPLICANT: Harley, Calvin B.  
; APPLICANT: Andrews, William H.  
; TITLE OF INVENTION: No. 6093809e1 Telomerase  
; NUMBER OF SEQUENCES: 225  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, 8th Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: United States of America  
; ZIP: 94111  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/851,843A  
; FILING DATE: 06-MAY-1997  
; CLASSIFICATION:  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: US 08/846,017  
; FILING DATE: 25-APR-1997  
; CLASSIFICATION:  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: US 08/844,419  
; FILING DATE: 18-APR-1997  
; CLASSIFICATION:  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: US 08/724,643  
; FILING DATE: 01-OCT-1996  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Apple, Randolph T.  
; REGISTRATION NUMBER: 36,429  
; REFERENCE/DOCKET NUMBER: 015389-002930US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 576-0200  
; TELEFAX: (415) 576-0300  
; INFORMATION FOR SEQ ID NO: 141:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 32 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide

;; FEATURE:  
;; NAME/KEY: Peptide  
;; LOCATION: 1..32  
;; OTHER INFORMATION: /note="motif 4(B') peptide from  
;; OTHER INFORMATION: Saccharomyces cerevisiae EST2p"  
US-08-851-843A-141

Query Match 68.4%; Score 26; DB 3; Length 32;  
Best Local Similarity 62.5%; Pred. No. 33;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLXDDLE 8  
::11111  
Db 20 LVDDLE 27

RESULT 38  
US-08-974-549A-259  
; Sequence 259, Application US/08974549A  
; Patent No. 6166178  
; GENERAL INFORMATION:  
; APPLICANT: Cech, Thomas R.  
; APPLICANT: Lingner, Joachim  
; APPLICANT: Nakamura, Toru  
; APPLICANT: Chapman, Karen B.  
; APPLICANT: Morin, Gregg B.  
; APPLICANT: Harley, Calvin B.  
; APPLICANT: Andrews, William H.  
; TITLE OF INVENTION: Human telomerase Catalytic Subunit  
; NUMBER OF SEQUENCES: 727  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/974,549A  
; FILING DATE: 19-NOV-1997  
; CLASSIFICATION: 536  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: US 08/724,643  
; FILING DATE: 01-OCT-1996  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: US 08/844,419  
; FILING DATE: 18-APR-1997  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: US 08/846,017  
; FILING DATE: 25-APR-1997  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: US 08/851,843  
; FILING DATE: 06-MAY-1997  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: US 08/854,050  
; FILING DATE: 09-MAY-1997  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: US 08/911,312  
; FILING DATE: 14-AUG-1997  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: US 08/912,951  
; FILING DATE: 14-AUG-1997  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: US 08/915,503  
; FILING DATE: 14-AUG-1997  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: WO PCT/US97/17618  
; FILING DATE: 01-OCT-1997

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/US97/17885  
FILING DATE: 01-OCT-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Apple, Randolph Ted  
REGISTRATION NUMBER: 36,429  
REFERENCE/DOCKET NUMBER: 015389-00261005  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 259:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 32 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Peptide  
LOCATION: 1..32  
OTHER INFORMATION: /note="motif 4(B') peptide from  
OTHER INFORMATION: Saccharomyces cerevisiae EST2p"  
US-08-974-549A-259

Query Match 68.4%; Score 26; DB 4; Length 32;  
Best Local Similarity 62.5%; Pred. No. 33;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Oy 1 VLXDDLE 8  
: : |||||  
Db 20 LVYDDLE 27

RESULT 39  
US-08-854-050-141  
Sequence 141, Application US/08854050  
Patent No. 6261836  
GENERAL INFORMATION:  
APPLICANT: Cech, Thomas R.  
APPLICANT: Lingner, Joachim  
APPLICANT: Nakamura, Toru  
APPLICANT: Chapman, Karen B.  
APPLICANT: Morlin, Gregg B.  
APPLICANT: Harley, Calvin  
APPLICANT: Andrews, William H.  
TITLE OF INVENTION: No. 6261836el Telomerase  
NUMBER OF SEQUENCES: 225  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, 8th Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: United States of America  
ZIP: 94111  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/854,050  
FILING DATE: 09-MAY-1997  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/851,843  
FILING DATE: 06-MAY-1997  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/846,017  
FILING DATE: 25-APR-1997  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/844,419

FILING DATE: 18-APR-1997  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/724,643  
FILING DATE: 01-OCT-1996  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Apple, Randolph T.  
REGISTRATION NUMBER: 36,429  
REFERENCE/DOCKET NUMBER: 015389-00293005  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 141:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 32 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Peptide  
LOCATION: 1..32  
OTHER INFORMATION: /note="motif 4(B') peptide from  
OTHER INFORMATION: Saccharomyces cerevisiae EST2p"  
US-08-854-050-141

Query Match 68.4%; Score 26; DB 4; Length 32;  
Best Local Similarity 62.5%; Pred. No. 33;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Oy 1 VLXDDLE 8  
: : |||||  
Db 20 LVYDDLE 27

RESULT 40  
US-09-430-323-141  
Sequence 141, Application US/09430323  
Patent No. 6309867  
GENERAL INFORMATION:  
APPLICANT: Cech, Thomas R.  
APPLICANT: Lingner, Joachim  
APPLICANT: Nakamura, Toru  
APPLICANT: Chapman, Karen B.  
APPLICANT: Morlin, Gregg B.  
APPLICANT: Harley, Calvin  
APPLICANT: Andrews, William H.  
TITLE OF INVENTION: No. 6309867el Telomerase  
NUMBER OF SEQUENCES: 225  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, 8th Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: United States of America  
ZIP: 94111  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/430,323  
FILING DATE: 29-OCT-1999  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/854,050  
FILING DATE: 09-MAY-1997  
APPLICATION NUMBER: US 08/851,843  
FILING DATE: 06-MAY-1997  
APPLICATION NUMBER: US 08/846,017  
FILING DATE: 25-APR-1997

APPLICATION NUMBER: US 08/844,419  
FILING DATE: 18-APR-1997  
APPLICATION NUMBER: US 08/724,643  
FILING DATE: 01-OCT-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Apple, Randolph T.  
REGISTRATION NUMBER: 36,429  
REFERENCE/DOCKET NUMBER: 015389-00293005  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 141:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 32 amino acids  
TYPE: amino acid  
STRANDEDNESS: <unknown>  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Peptide  
LOCATION: 1..32  
OTHER INFORMATION: /note="motif 4(B') peptide from  
Saccharomyces cerevisiae EST2p"  
SEQUENCE DESCRIPTION: SEQ ID NO: 141:  
US-09-430-323-141  
Query Match 68.4%; Score 26; DB 4; Length 32;  
Best Local Similarity 62.5%; Pred. No. 33;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
QY 1 VLXDDLE 8  
DB 20 LVYDDLE 27  
RESULT 41  
US-08-974-549A-46  
Sequence 46, Application US/08974549A  
Patent No. 6166178  
GENERAL INFORMATION:  
APPLICANT: Cecch, Thomas R.  
APPLICANT: Lingner, Joachim  
APPLICANT: Nakamura, Toru  
APPLICANT: Chapman, Karen B.  
APPLICANT: Morin, Gregg B.  
APPLICANT: Harley, Calvin B.  
APPLICANT: Andrews, William H.  
TITLE OF INVENTION: Human Telomerase Catalytic Subunit  
NUMBER OF SEQUENCES: 727  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/974,549A  
FILING DATE: 19-NOV-1997  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/724,643  
FILING DATE: 01-OCT-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/844,419  
FILING DATE: 18-APR-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/846,017

FILING DATE: 25-APR-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/851,843  
FILING DATE: 06-MAY-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/854,050  
FILING DATE: 09-MAY-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/911,312  
FILING DATE: 14-AUG-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/912,951  
FILING DATE: 14-AUG-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/915,503  
FILING DATE: 14-AUG-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/US97/17618  
FILING DATE: 01-OCT-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/US97/17885  
FILING DATE: 01-OCT-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Apple, Randolph T.  
REGISTRATION NUMBER: 36,429  
REFERENCE/DOCKET NUMBER: 015389-00261005  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 46:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 35 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Peptide  
LOCATION: 1..35  
OTHER INFORMATION: /note="motif B' peptide from  
Saccharomyces cerevisiae EST2p"  
US-08-974-549A-46  
Query Match 68.4%; Score 26; DB 4; Length 35;  
Best Local Similarity 62.5%; Pred. No. 37;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
QY 1 VLXDDLE 8  
DB 22 LVYDDLE 29  
RESULT 42  
US-08-851-843A-124  
Sequence 124, Application US/08851843A  
Patent No. 6093809  
GENERAL INFORMATION:  
APPLICANT: Cecch, Thomas R.  
APPLICANT: Lingner, Joachim  
APPLICANT: Nakamura, Toru  
APPLICANT: Chapman, Karen B.  
APPLICANT: Morin, Gregg B.  
APPLICANT: Harley, Calvin B.  
APPLICANT: Andrews, William H.  
TITLE OF INVENTION: No. 6093809e1 Telomerase  
NUMBER OF SEQUENCES: 225  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, 8th Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: United States of America  
ZIP: 94111

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/851,843A  
FILING DATE: 06-MAY-1997  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/846,017  
FILING DATE: 25-APR-1997  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/844,419  
FILING DATE: 18-APR-1997  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/724,643  
FILING DATE: 01-OCT-1996  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Apple, Randolph T.  
REGISTRATION NUMBER: 36,429  
REFERENCE/DOCKET NUMBER: 015389-00293005  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 124:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 49 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-851-843A-124

Query Match 68.4%; Score 26; DB 3; Length 49;  
Best Local Similarity 62.5%; Pred. No. 53;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLXDPLE 8  
Db 16 LVYDPLE 23

RESULT 43  
US-08-974-549A-244  
Sequence 244, Application US/08974549A  
Patent No. 6166178  
GENERAL INFORMATION:  
APPLICANT: Cecch, Thomas R.  
APPLICANT: Lingner, Joachim  
APPLICANT: Nakamura, Toru  
APPLICANT: Chapman, Karen B.  
APPLICANT: Morin, Gregg B.  
APPLICANT: Harley, Calvin B.  
APPLICANT: Andrews, William H.  
TITLE OF INVENTION: Human Telomerase Catalytic Subunit  
NUMBER OF SEQUENCES: 727  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/974,549A  
FILING DATE: 19-NOV-1997  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/724,643  
FILING DATE: 01-OCT-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/844,419  
FILING DATE: 18-APR-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/846,017  
FILING DATE: 25-APR-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/851,843  
FILING DATE: 06-MAY-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/854,050  
FILING DATE: 09-MAY-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/911,312  
FILING DATE: 14-AUG-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/912,951  
FILING DATE: 14-AUG-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/915,503  
FILING DATE: 14-AUG-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/US97/17618  
FILING DATE: 01-OCT-1997  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/US97/17885  
FILING DATE: 01-OCT-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Apple, Randolph Ted  
REGISTRATION NUMBER: 36,429  
REFERENCE/DOCKET NUMBER: 015389-00261005  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 244:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 49 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-974-549A-244

Query Match 68.4%; Score 26; DB 4; Length 49;  
Best Local Similarity 62.5%; Pred. No. 53;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLXDPLE 8  
Db 16 LVYDPLE 23

RESULT 44  
US-08-854-050-124  
Sequence 124, Application US/08854050  
Patent No. 6261836  
GENERAL INFORMATION:  
APPLICANT: Cecch, Thomas R.  
APPLICANT: Lingner, Joachim  
APPLICANT: Nakamura, Toru  
APPLICANT: Chapman, Karen B.  
APPLICANT: Morin, Gregg B.  
APPLICANT: Harley, Calvin B.  
APPLICANT: Andrews, William H.  
TITLE OF INVENTION: No. 6261836el Telomerase  
NUMBER OF SEQUENCES: 225  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, 8th Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: United States of America  
ZIP: 94111

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/854,050  
FILING DATE: 09-MAY-1997  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/851,843  
FILING DATE: 06-MAY-1997  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/846,017  
FILING DATE: 25-APR-1997  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/844,419  
FILING DATE: 18-APR-1997  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/724,643  
FILING DATE: 01-OCT-1996  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Apple, Randolph T.  
REGISTRATION NUMBER: 36,429  
REFERENCE/DOCKET NUMBER: 015389-00293005  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 124:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 49 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-854-050-124

Query Match 68.4%; Score 26; DB 4; Length 49;  
Best Local Similarity 62.5%; Pred. No. 53;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 LVXDLE 8  
DB 16 LVYDLE 23

RESULT 45  
US-09-430-323-124  
Sequence 124, Application US/09430323  
Patent No. 6309867  
GENERAL INFORMATION:  
APPLICANT: Cecch, Thomas R.  
Lingner, Joachim  
Nakamura, Toru  
Chapman, Karen B.  
Morin, Gregg B.  
Harley, Calvin  
Andrews, William H.  
TITLE OF INVENTION: No. 6309867e1 Telomerase  
NUMBER OF SEQUENCES: 225  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, 8th Floor

CITY: San Francisco  
STATE: California  
COUNTRY: United States of America  
ZIP: 94111

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/430,323  
FILING DATE: 29-OCT-1999  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/854,050  
FILING DATE: 09-MAY-1997  
APPLICATION NUMBER: US 08/851,843  
FILING DATE: 06-MAY-1997  
APPLICATION NUMBER: US 08/846,017  
FILING DATE: 25-APR-1997  
APPLICATION NUMBER: US 08/844,419  
FILING DATE: 18-APR-1997  
APPLICATION NUMBER: US 08/724,643  
FILING DATE: 01-OCT-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Apple, Randolph T.  
REGISTRATION NUMBER: 36,429  
REFERENCE/DOCKET NUMBER: 015389-00293005  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 124:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 49 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 124:  
US-09-430-323-124

Query Match 68.4%; Score 26; DB 4; Length 49;  
Best Local Similarity 62.5%; Pred. No. 53;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 LVXDLE 8  
DB 16 LVYDLE 23

RESULT 46  
US-08-851-843A-17  
Sequence 17, Application US/08851843A  
Patent No. 6093809  
GENERAL INFORMATION:  
APPLICANT: Cecch, Thomas R.  
Lingner, Joachim  
Nakamura, Toru  
Chapman, Karen B.  
Morin, Gregg B.  
Harley, Calvin  
Andrews, William H.  
TITLE OF INVENTION: No. 6093809e1 Telomerase  
NUMBER OF SEQUENCES: 225  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, 8th Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: United States of America  
ZIP: 94111  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/851,843A  
FILING DATE: 06-MAY-1997  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/846,017  
FILING DATE: 25-APR-1997  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/844,419  
FILING DATE: 18-APR-1997  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/724,643  
FILING DATE: 01-OCT-1996  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Apple, Randolph T.  
REGISTRATION NUMBER: 36,429  
REFERENCE/DOCKET NUMBER: 015389-002930US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 69 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: not relevant  
MOLECULE TYPE: peptide  
US-08-851-843A-17

Query Match 68.4%; Score 26; DB 3; Length 69;  
Best Local Similarity 62.5%; Pred. No. 78;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 LVXDLE 8  
: : |||||  
Db 56 LVYDDLE 63

RESULT 47  
US-08-854-050-17  
Sequence 17, Application US/08854050  
Patent No. 6261836  
GENERAL INFORMATION:  
APPLICANT: Cech, Thomas R.  
APPLICANT: Lingner, Joachim  
APPLICANT: Nakamura, Toru  
APPLICANT: Chapman, Karen B.  
APPLICANT: Morin, Gregg B.  
APPLICANT: Harley, Calvin  
APPLICANT: Andrews, William H.  
TITLE OF INVENTION: No. 6261836el Telomerase  
NUMBER OF SEQUENCES: 225  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, 8th Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: United States of America  
ZIP: 94111  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/854,050  
FILING DATE: 09-MAY-1997

CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/851,843  
FILING DATE: 06-MAY-1997  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/846,017  
FILING DATE: 25-APR-1997  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/844,419  
FILING DATE: 18-APR-1997  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/724,643  
FILING DATE: 01-OCT-1996  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Apple, Randolph T.  
REGISTRATION NUMBER: 36,429  
REFERENCE/DOCKET NUMBER: 015389-002930US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 69 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: not relevant  
MOLECULE TYPE: peptide  
US-08-854-050-17

Query Match 68.4%; Score 26; DB 4; Length 69;  
Best Local Similarity 62.5%; Pred. No. 78;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 LVXDLE 8  
: : |||||  
Db 56 LVYDDLE 63

RESULT 48  
US-09-430-323-17  
Sequence 17, Application US/09430323  
Patent No. 6309867  
GENERAL INFORMATION:  
APPLICANT: Cech, Thomas R.  
APPLICANT: Lingner, Joachim  
APPLICANT: Nakamura, Toru  
APPLICANT: Chapman, Karen B.  
APPLICANT: Morin, Gregg B.  
APPLICANT: Harley, Calvin  
APPLICANT: Andrews, William H.  
TITLE OF INVENTION: No. 6309867el Telomerase  
NUMBER OF SEQUENCES: 225  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, 8th Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: United States of America  
ZIP: 94111  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/430,323  
FILING DATE: 29-OCT-1999  
CLASSIFICATION: <unknown>  
PRIOR APPLICATION DATA:



```

; APPLICATION NUMBER: US 08/854,050
; FILING DATE: 09-MAY-1997
; APPLICATION NUMBER: US 08/851,843
; FILING DATE: 06-MAY-1997
; APPLICATION NUMBER: US 08/846,017
; FILING DATE: 25-APR-1997
; APPLICATION NUMBER: US 08/844,419
; FILING DATE: 18-APR-1997
; APPLICATION NUMBER: US 08/724,643
; FILING DATE: 01-OCT-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Apple, Randolph T.
; REGISTRATION NUMBER: 36,429
; REFERENCE/DOCKET NUMBER: 015389-002930US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 69 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 17:
US-09-430-323-17

```

```

Query Match      68.4%; Score 26; DB 4; Length 69;
Best Local Similarity 62.5%; Pred. No. 78;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 LXDLDLE 8
Db 56 LVYDLDLE 63

```

```

RESULT 49
US-09-153-599A-5
; Sequence 5, Application US/09153599A
; Patent No. 6420177
; GENERAL INFORMATION:
; APPLICANT: Weber, J. Mark
; APPLICANT: Luu, B. Minh
; TITLE OF INVENTION: Method for Strain Improvement of
; TITLE OF INVENTION: Erythromycin Producing Bacterium
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rokey, Milamow & Katz, Ltd.
; STREET: 180 N. Stetson Avenue, 2 Prudential Plaza
; CITY: Chicago
; STATE: Illinois
; COUNTRY: U.S.A.
; ZIP: 60601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/153,599A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mueller, Lisa V.
; REGISTRATION NUMBER: 38,978
; REFERENCE/DOCKET NUMBER: FER2159P004IUS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-616-5460
; TELEFAX: 312-616-5400
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 241 amino acids
; TYPE: amino acid

```

```

; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-153-599A-5

```

```

Query Match      68.4%; Score 26; DB 4; Length 241;
Best Local Similarity 62.5%; Pred. No. 3.1e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 2 LXDLDLEA 9
Db 150 LVYDLDLES 157

```

```

RESULT 50
US-09-134-001C-5669
; Sequence 5669, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 5669
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5669

```

```

Query Match      68.4%; Score 26; DB 4; Length 291;
Best Local Similarity 71.4%; Pred. No. 3.9e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 2 LXDLDLE 8
Db 117 LVYDLDLE 123

```

```

Search completed: December 27, 2002, 14:41:51
Job time : 27 secs

```



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 27, 2002, 14:42:33 ; Search time 10 Seconds  
(Without alignments)  
15.583 Million cell updates/sec

Title: US-09-489-760A-1  
Perfect score: 38  
Sequence: 1 VLXDDLEA 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapept 0.5

Searched: 109717 seqs, 1731436 residues

Total number of hits satisfying chosen parameters: 109717

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 1000 summaries

Database :

Published\_Applications\_AA:\*  
1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*  
2: /cgn2\_6/ptodata/2/pubpaa/PC1\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/2/pubpaa/PC105\_PUBCOMB.pep:\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep:\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep:\*  
12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep:\*  
13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:\*  
14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	30	78.9	94	10	US-09-764-847-543
2	30	78.9	1410	9	US-10-014-717-3
3	29	76.3	594	9	US-09-989-920-200
4	29	76.3	594	9	US-09-989-920-249
5	28	73.7	293	10	US-09-934-899-2
6	28	73.7	293	10	US-09-934-868-22
7	28	73.7	296	10	US-09-815-242-13353
8	28	73.7	296	10	US-09-815-242-13677
9	27	71.1	63	10	US-09-864-761-11433
10	27	71.1	140	12	US-10-052-586-380
11	27	71.1	293	10	US-09-815-242-10630
12	27	71.1	419	10	US-09-815-915-11
13	27	71.1	602	10	US-09-952-013A-6
14	27	71.1	664	10	US-09-735-927-2
15	27	71.1	664	10	US-09-735-927-4
16	27	71.1	664	12	US-10-034-843-2
17	27	71.1	1172	9	US-09-712-363-176
18	26	68.4	32	9	US-09-843-676-141
19	26	68.4	32	9	US-09-766-253-141

20	26	68.4	37	10	US-09-823-266-19	Sequence 19, Appl
21	26	68.4	45	10	US-09-864-761-36279	Sequence 36279, A
22	26	68.4	49	9	US-09-843-676-124	Sequence 124, App
23	26	68.4	49	9	US-09-766-253-124	Sequence 124, App
24	26	68.4	69	9	US-09-843-676-17	Sequence 17, Appl
25	26	68.4	69	9	US-09-766-253-17	Sequence 17, Appl
26	26	68.4	148	10	US-09-867-550-10	Sequence 10, Appl
27	26	68.4	151	10	US-09-925-302-599	Sequence 599, App
28	26	68.4	130	10	US-09-925-297-793	Sequence 793, App
29	26	68.4	333	12	US-10-013-056-2	Sequence 2, Appli
30	26	68.4	333	12	US-10-013-056-2	Sequence 2, Appli
31	26	68.4	448	10	US-09-754-446-4	Sequence 4, Appli
32	26	68.4	452	9	US-09-905-291A-255	Sequence 255, App
33	26	68.4	452	9	US-09-902-853-255	Sequence 255, App
34	26	68.4	452	10	US-09-909-320-255	Sequence 255, App
35	26	68.4	452	10	US-09-909-088B-255	Sequence 255, App
36	26	68.4	457	10	US-09-754-446-2	Sequence 2, Appli
37	26	68.4	457	10	US-09-815-242-13478	Sequence 13478, A
38	26	68.4	568	12	US-10-071-751-30	Sequence 30, Appl
39	26	68.4	568	12	US-10-071-751-30	Sequence 30, Appl
40	26	68.4	660	12	US-10-115-178-1	Sequence 1, Appli
41	26	68.4	709	10	US-09-767-870-8	Sequence 8, Appli
42	26	68.4	736	9	US-10-060-230-23	Sequence 23, Appl
43	26	68.4	736	9	US-10-060-230-24	Sequence 24, Appl
44	26	68.4	780	10	US-09-770-689A-5	Sequence 5, Appli
45	26	68.4	803	10	US-09-770-689A-2	Sequence 2, Appli
46	26	68.4	829	10	US-09-815-242-11216	Sequence 11216, A
47	26	68.4	868	9	US-10-041-007-4	Sequence 4, Appli
48	26	68.4	868	10	US-09-887-586A-56	Sequence 56, Appl
49	26	68.4	868	10	US-09-903-012-56	Sequence 56, Appl
50	26	68.4	881	10	US-09-816-860A-2	Sequence 2, Appli
51	26	68.4	884	9	US-09-843-676-55	Sequence 55, Appl
52	26	68.4	884	9	US-09-766-253-55	Sequence 55, Appl
53	26	68.4	1182	10	US-09-815-242-11828	Sequence 11828, A
54	26	68.4	5405	9	US-10-025-380-1116	Sequence 1116, Ap
55	26	68.4	5405	10	US-09-922-217-1116	Sequence 1116, Ap
56	26	68.4	43	10	US-09-864-761-35681	Sequence 35681, A
57	26	68.4	46	10	US-09-864-761-39680	Sequence 39680, A
58	26	68.4	54	10	US-09-764-887-302	Sequence 302, App
59	26	68.4	57	10	US-09-976-280A-25	Sequence 25, Appl
60	26	68.4	62	10	US-09-864-761-42226	Sequence 42226, A
61	26	68.4	62	10	US-09-925-300-1280	Sequence 1280, Ap
62	26	68.4	107	10	US-09-867-550-676	Sequence 676, App
63	26	68.4	117	10	US-09-765-272-128	Sequence 128, App
64	26	68.4	121	10	US-09-741-666-447	Sequence 447, App
65	26	68.4	144	10	US-09-515-806-16	Sequence 16, Appl
66	26	68.4	181	10	US-09-765-272-62	Sequence 62, Appl
67	26	68.4	205	10	US-09-841-133-564	Sequence 564, App
68	26	68.4	230	10	US-09-216-393-31	Sequence 31, Appl
69	26	68.4	242	10	US-09-815-242-11694	Sequence 11694, A
70	26	68.4	250	10	US-09-833-017-6	Sequence 6, Appli
71	26	68.4	252	10	US-09-987-446-2	Sequence 2, Appli
72	26	68.4	275	10	US-09-815-242-13441	Sequence 13441, A
73	26	68.4	280	9	US-09-764-868-1125	Sequence 1125, Ap
74	26	68.4	283	9	US-09-764-868-712	Sequence 712, App
75	26	68.4	289	9	US-10-023-446-4	Sequence 4, Appli
76	26	68.4	315	10	US-09-815-242-10142	Sequence 10142, A
77	26	68.4	319	10	US-09-815-242-5009	Sequence 5009, Ap
78	26	68.4	321	10	US-09-815-242-13978	Sequence 13978, A
79	26	68.4	322	10	US-09-815-242-12031	Sequence 12031, A
80	26	68.4	323	10	US-09-815-242-10321	Sequence 10321, A
81	26	68.4	323	10	US-09-815-242-13734	Sequence 13734, A
82	26	68.4	327	10	US-09-815-242-10896	Sequence 10896, A
83	26	68.4	340	10	US-09-925-302-655	Sequence 655, App
84	26	68.4	343	10	US-09-971-798-27	Sequence 27, Appl
85	26	68.4	368	10	US-09-815-242-5636	Sequence 5636, Ap
86	26	68.4	368	10	US-09-815-242-12393	Sequence 12393, A
87	26	68.4	376	10	US-09-815-242-13934	Sequence 13934, A
88	26	68.4	384	10	US-09-801-876B-6	Sequence 6, Appli
89	26	68.4	385	10	US-09-925-302-624	Sequence 624, App
90	26	68.4	401	10	US-09-729-674-2	Sequence 2, Appli
91	26	68.4	427	10	US-09-912-020-324	Sequence 324, App
92	26	68.4	461	10	US-09-764-870-365	Sequence 365, App

93	25	65.8	464	10	US-09-815-242-10647	Sequence 10647, A	166	24	63.2	318	10	US-09-815-242-11531	Sequence 11531, A
94	25	65.8	467	10	US-09-815-242-4997	Sequence 4997, Ap	167	24	63.2	330	10	US-09-815-242-11914	Sequence 11914, A
95	25	65.8	471	10	US-09-815-242-10682	Sequence 10682, A	168	24	63.2	339	10	US-09-815-242-13235	Sequence 13235, A
96	25	65.8	478	10	US-09-815-242-11674	Sequence 11674, A	169	24	63.2	339	10	US-09-815-242-13635	Sequence 13635, A
97	25	65.8	508	10	US-09-826-752-10	Sequence 10, Appl	170	24	63.2	343	10	US-09-794-960-4	Sequence 4, Appl
98	25	65.8	511	10	US-09-815-242-10737	Sequence 10737, A	171	24	63.2	355	10	US-09-789-482-4	Sequence 4, Appl
99	25	65.8	591	9	US-09-201-936-42	Sequence 42, Appl	172	24	63.2	355	10	US-09-789-486-4	Sequence 4, Appl
100	25	65.8	612	10	US-09-974-592-14	Sequence 14, Appl	173	24	63.2	364	10	US-09-815-242-13652	Sequence 13652, A
101	25	65.8	621	10	US-09-815-242-11303	Sequence 11303, A	174	24	63.2	369	10	US-09-838-965-3	Sequence 3, Appl
102	25	65.8	630	10	US-09-815-242-12117	Sequence 12117, A	175	24	63.2	371	10	US-09-815-242-5667	Sequence 5667, Ap
103	25	65.8	645	10	US-09-862-027-15	Sequence 15, Appl	176	24	63.2	374	9	US-09-974-298-95	Sequence 95, Appl
104	25	65.8	645	10	US-09-862-027-16	Sequence 16, Appl	177	24	63.2	375	10	US-09-783-320-18	Sequence 18, Appl
105	25	65.8	650	10	US-09-854-875A-1	Sequence 1, Appl	178	24	63.2	380	10	US-09-815-242-13177	Sequence 13177, A
106	25	65.8	650	12	US-10-056-253-1	Sequence 1, Appl	179	24	63.2	380	10	US-09-784-810A-14	Sequence 14, Appl
107	25	65.8	672	10	US-09-858-754-2	Sequence 2, Appl	180	24	63.2	390	9	US-09-971-536-379	Sequence 53, Appl
108	25	65.8	673	10	US-09-841-835-8	Sequence 8, Appl	181	24	63.2	391	10	US-09-815-242-5379	Sequence 5379, Ap
109	25	65.8	705	10	US-09-815-242-5334	Sequence 5334, Ap	182	24	63.2	393	10	US-09-815-242-12285	Sequence 12285, A
110	25	65.8	713	10	US-09-815-242-12322	Sequence 12322, A	183	24	63.2	394	10	US-09-205-448-2	Sequence 2, Appl
111	25	65.8	734	10	US-09-946-931-2	Sequence 2, Appl	184	24	63.2	400	10	US-09-815-242-12208	Sequence 12208, A
112	25	65.8	734	10	US-09-946-931-4	Sequence 4, Appl	185	24	63.2	402	10	US-09-888-243-14	Sequence 14, Appl
113	25	65.8	819	10	US-09-833-790-425	Sequence 425, App	186	24	63.2	409	9	US-10-157-855-22	Sequence 22, Appl
114	25	65.8	840	10	US-09-815-242-5198	Sequence 5199, Ap	187	24	63.2	416	10	US-09-801-368-54	Sequence 54, Appl
115	25	65.8	879	9	US-10-108-605-217	Sequence 217, App	188	24	63.2	457	10	US-09-815-242-12670	Sequence 12670, A
116	25	65.8	949	10	US-09-841-835-10	Sequence 10, Appl	189	24	63.2	473	10	US-09-783-320-22	Sequence 22, Appl
117	25	65.8	981	10	US-09-991-258-13	Sequence 13, Appl	190	24	63.2	513	9	US-10-020-038-8	Sequence 8, Appl
118	25	65.8	1034	10	US-09-858-754-5	Sequence 5, Appl	191	24	63.2	520	10	US-09-783-320-30	Sequence 30, Appl
119	25	65.8	1159	10	US-09-735-995-2	Sequence 2, Appl	192	24	63.2	523	10	US-09-815-242-5723	Sequence 5723, Ap
120	25	65.8	1159	10	US-09-735-995-4	Sequence 4, Appl	193	24	63.2	525	10	US-09-815-242-12651	Sequence 12651, A
121	25	65.8	1159	10	US-09-119-895-10	Sequence 10, Appl	194	24	63.2	534	10	US-09-764-864-871	Sequence 871, App
122	25	65.8	1263	10	US-09-971-309-6	Sequence 6, Appl	195	24	63.2	560	10	US-09-783-320-38	Sequence 38, Appl
123	25	65.8	1302	12	US-10-000-864-2	Sequence 2, Appl	196	24	63.2	576	10	US-09-783-320-46	Sequence 46, Appl
124	25	65.8	1327	10	US-09-841-835-2	Sequence 2, Appl	197	24	63.2	576	10	US-09-994-288-5	Sequence 5, Appl
125	25	65.8	1464	9	US-09-842-777-10	Sequence 10, Appl	198	24	63.2	591	10	US-09-815-242-5837	Sequence 5837, Ap
126	25	65.8	1493	10	US-09-858-754-3	Sequence 3, Appl	199	24	63.2	594	12	US-10-062-254-952	Sequence 952, App
127	25	65.8	1493	10	US-09-858-754-4	Sequence 4, Appl	200	24	63.2	599	10	US-09-952-013A-4	Sequence 4, Appl
128	25	65.8	1493	12	US-10-000-864-8	Sequence 8, Appl	201	24	63.2	610	10	US-09-853-118-2	Sequence 118-2, A
129	25	65.8	2472	10	US-09-815-242-5064	Sequence 5064, Ap	202	24	63.2	617	10	US-09-815-242-5762	Sequence 5762, Ap
130	24	63.2	42	10	US-09-864-761-37105	Sequence 37105, A	203	24	63.2	624	10	US-09-815-242-5640	Sequence 5640, Ap
131	24	63.2	46	10	US-09-864-761-39461	Sequence 39461, A	204	24	63.2	624	10	US-09-815-242-12704	Sequence 12704, A
132	24	63.2	50	10	US-09-864-761-39753	Sequence 39753, A	205	24	63.2	625	10	US-09-881-752A-144	Sequence 144, App
133	24	63.2	54	10	US-09-764-847-538	Sequence 538, App	206	24	63.2	628	10	US-09-815-242-13002	Sequence 13002, A
134	24	63.2	86	10	US-09-939-980-294	Sequence 294, App	207	24	63.2	701	10	US-09-768-877-26	Sequence 26, Appl
135	24	63.2	89	10	US-09-815-242-12185	Sequence 12185, A	208	24	63.2	701	10	US-09-815-242-13002	Sequence 13002, A
136	24	63.2	89	10	US-09-815-242-12143	Sequence 12743, A	209	24	63.2	714	10	US-09-768-877-22	Sequence 22, Appl
137	24	63.2	108	9	US-09-864-921-90	Sequence 90, Appl	210	24	63.2	738	9	US-09-712-363-238	Sequence 238, App
138	24	63.2	108	9	US-09-864-921-91	Sequence 21, Appl	211	24	63.2	740	10	US-09-815-242-10376	Sequence 10376, A
139	24	63.2	119	10	US-09-989-903-21	Sequence 21, Appl	212	24	63.2	743	10	US-09-815-242-11006	Sequence 11006, A
140	24	63.2	139	10	US-09-864-761-34885	Sequence 34885, A	213	24	63.2	749	10	US-09-771-161A-261	Sequence 261, App
141	24	63.2	139	10	US-09-893-737-264	Sequence 264, Appl	214	24	63.2	782	10	US-09-841-739-9	Sequence 9, Appl
142	24	63.2	142	12	US-10-004-832-2	Sequence 2, Appl	215	24	63.2	900	10	US-09-815-242-13381	Sequence 13381, A
143	24	63.2	143	10	US-09-925-302-749	Sequence 749, App	216	24	63.2	900	10	US-09-815-242-13663	Sequence 13663, A
144	24	63.2	147	10	US-09-925-302-642	Sequence 642, App	217	24	63.2	905	9	US-10-114-893-127	Sequence 127, App
145	24	63.2	148	12	US-10-052-586-292	Sequence 292, App	218	24	63.2	912	9	US-09-865-960-2	Sequence 2, Appl
146	24	63.2	162	10	US-09-734-569-16	Sequence 16, App	219	24	63.2	915	10	US-09-880-192-59	Sequence 59, Appl
147	24	63.2	208	10	US-09-764-870-497	Sequence 497, App	220	24	63.2	969	9	US-10-023-437-33	Sequence 33, Appl
148	24	63.2	210	10	US-09-864-761-33812	Sequence 33812, A	221	24	63.2	972	10	US-09-944-607-10	Sequence 10, Appl
149	24	63.2	226	10	US-09-815-242-5448	Sequence 5448, Ap	222	24	63.2	999	9	US-09-895-913A-226	Sequence 226, App
150	24	63.2	223	10	US-10-047-260-16	Sequence 16, Appl	223	24	63.2	1134	9	US-09-836-392-15	Sequence 15, Appl
151	24	63.2	237	10	US-09-765-272-136	Sequence 136, App	224	24	63.2	1173	10	US-09-815-242-11935	Sequence 11935, A
152	24	63.2	254	10	US-09-815-242-11331	Sequence 11331, A	225	24	63.2	1192	10	US-09-815-242-10903	Sequence 10903, A
153	24	63.2	258	9	US-10-023-437-29	Sequence 29, Appl	226	24	63.2	1237	10	US-09-862-027-78	Sequence 78, Appl
154	24	63.2	275	10	US-09-815-242-10842	Sequence 10842, A	227	24	63.2	1286	9	US-10-017-216-7	Sequence 7, Appl
155	24	63.2	282	10	US-09-815-242-5517	Sequence 5517, Ap	228	24	63.2	1294	12	US-10-017-223-2	Sequence 2, Appl
156	24	63.2	290	10	US-09-764-864-1323	Sequence 1323, Ap	229	24	63.2	1352	9	US-09-736-968A-9	Sequence 9, Appl
157	24	63.2	295	10	US-09-764-870-364	Sequence 364, App	230	24	63.2	1352	10	US-09-736-969A-9	Sequence 9, Appl
158	24	63.2	296	10	US-09-783-320-8	Sequence 8, Appl	231	24	63.2	1352	10	US-09-736-969-9	Sequence 9, Appl
159	24	63.2	303	10	US-09-815-242-12132	Sequence 12132, A	232	24	63.2	1353	10	US-09-751-1008-2	Sequence 2, Appl
160	24	63.2	303	10	US-09-925-297-740	Sequence 740, App	233	24	63.2	1353	10	US-09-751-1008-99	Sequence 99, Appl
161	24	63.2	310	10	US-09-734-569-148	Sequence 148, App	234	24	63.2	1353	12	US-10-071-223-3	Sequence 3, Appl
162	24	63.2	313	10	US-09-815-242-12040	Sequence 12040, A	235	24	63.2	1403	8	US-08-913-322-22	Sequence 22, Appl
163	24	63.2	315	10	US-09-815-242-11261	Sequence 11261, A	236	24	63.2	1534	9	US-08-913-322-24	Sequence 24, Appl
164	24	63.2	318	10	US-09-783-320-12	Sequence 12, Appl	237	24	63.2	1534	9	US-09-736-968A-10	Sequence 10, Appl
165	24	63.2	318	10	US-09-815-242-11362	Sequence 11362, A	238	24	63.2	1534	10	US-09-736-969A-10	Sequence 10, Appl

229	24	63.2	1534	10	US-09-736-960-10	Sequence 10, Appl	312	23	60.5	270	10	US-09-925-302-677	Sequence 677, App
240	24	63.2	1597	9	US-10-017-216-6	Sequence 6, Appl	313	23	60.5	271	9	US-09-974-298-6	Sequence 6, Appl
241	24	63.2	1641	9	US-10-017-216-5	Sequence 5, Appl	314	23	60.5	271	9	US-09-770-528-11	Sequence 11, Appl
242	24	63.2	1770	10	US-09-801-368-298	Sequence 298, App	315	23	60.5	271	10	US-09-005-243-52	Sequence 52, Appl
243	24	63.2	1958	12	US-10-028-946-4	Sequence 4, Appl	316	23	60.5	271	10	US-09-224-683-52	Sequence 52, Appl
244	24	63.2	1980	9	US-09-736-968A-108	Sequence 108, App	317	23	60.5	271	10	US-09-854-280-25	Sequence 25, Appl
245	24	63.2	1980	10	US-09-736-968A-94	Sequence 94, Appl	318	23	60.5	271	10	US-09-840-707A-1	Sequence 1, Appl
246	24	63.2	1980	10	US-09-736-960-91	Sequence 91, Appl	319	23	60.5	271	10	US-09-854-208-25	Sequence 25, Appl
247	24	63.2	2053	9	US-10-017-216-2	Sequence 2, Appl	320	23	60.5	271	10	US-09-919-172-7	Sequence 7, Appl
248	24	63.2	2054	12	US-10-028-946-2	Sequence 2, Appl	321	23	60.5	273	10	US-09-005-243-48	Sequence 48, Appl
249	24	63.2	2055	9	US-10-017-216-4	Sequence 4, Appl	322	23	60.5	273	10	US-09-005-243-49	Sequence 49, Appl
250	24	63.2	3503	9	US-10-108-605-237	Sequence 237, App	323	23	60.5	273	10	US-09-005-243-50	Sequence 50, Appl
251	23	60.5	9	10	US-09-834-765-132	Sequence 132, App	324	23	60.5	273	10	US-09-005-243-53	Sequence 53, Appl
252	23	60.5	9	10	US-09-834-765-349	Sequence 349, App	325	23	60.5	273	10	US-09-005-243-61	Sequence 61, Appl
253	23	60.5	10	10	US-09-834-765-195	Sequence 195, App	326	23	60.5	273	10	US-09-224-683-48	Sequence 48, Appl
254	23	60.5	22	1	US-08-841-636A-24	Sequence 24, Appl	327	23	60.5	273	10	US-09-224-683-49	Sequence 49, Appl
255	23	60.5	25	10	US-09-758-140-16	Sequence 16, Appl	328	23	60.5	273	10	US-09-224-683-50	Sequence 50, Appl
256	23	60.5	25	10	US-09-972-599A-16	Sequence 16, Appl	329	23	60.5	273	10	US-09-224-683-53	Sequence 53, Appl
257	23	60.5	31	10	US-09-864-761-3355	Sequence 3355, A	330	23	60.5	273	10	US-09-224-683-61	Sequence 61, Appl
258	23	60.5	34	10	US-09-864-761-4073A	Sequence 4073A, A	331	23	60.5	274	10	US-09-005-243-51	Sequence 51, Appl
259	23	60.5	37	10	US-09-864-761-37270	Sequence 37270, A	332	23	60.5	274	10	US-09-224-683-51	Sequence 51, Appl
260	23	60.5	61	10	US-09-829-481-4	Sequence 43297, A	333	23	60.5	278	10	US-09-764-853-559	Sequence 559, App
261	23	60.5	67	10	US-09-864-761-43297	Sequence 4, Appl	334	23	60.5	286	10	US-09-925-300-1296	Sequence 1296, App
262	23	60.5	66	10	US-09-758-140-20	Sequence 20, Appl	335	23	60.5	287	10	US-09-925-637-64	Sequence 64, Appl
263	23	60.5	66	10	US-09-972-599A-20	Sequence 20, Appl	336	23	60.5	288	10	US-09-815-242-5239	Sequence 5239, App
264	23	60.5	66	10	US-09-972-599A-22	Sequence 22, Appl	337	23	60.5	293	10	US-09-815-242-12583	Sequence 12583, A
265	23	60.5	66	10	US-09-864-761-42185	Sequence 42185, A	338	23	60.5	296	9	US-10-106-534-4	Sequence 4, Appl
266	23	60.5	70	10	US-09-864-761-42173	Sequence 46173, A	339	23	60.5	304	10	US-09-901-884-1	Sequence 1, Appl
267	23	60.5	76	10	US-09-864-761-36556	Sequence 36556, A	340	23	60.5	306	10	US-09-815-242-11035	Sequence 11035, A
268	23	60.5	86	10	US-09-864-761-37632	Sequence 37632, A	341	23	60.5	307	10	US-09-815-242-11306	Sequence 11306, A
269	23	60.5	91	10	US-09-864-761-46054	Sequence 46054, A	342	23	60.5	312	10	US-09-746-491-26	Sequence 26, Appl
270	23	60.5	92	10	US-09-864-761-47663	Sequence 47663, A	343	23	60.5	322	10	US-09-815-242-11634	Sequence 11634, A
271	23	60.5	94	10	US-09-864-761-45080	Sequence 45080, A	344	23	60.5	324	10	US-09-866-562-40	Sequence 40, Appl
272	23	60.5	94	10	US-09-864-761-48815	Sequence 48815, A	345	23	60.5	326	12	US-10-062-994-28	Sequence 28, Appl
273	23	60.5	101	10	US-09-730-525-37	Sequence 37, Appl	346	23	60.5	326	12	US-10-062-994-28	Sequence 28, Appl
274	23	60.5	101	10	US-09-730-917-37	Sequence 37, Appl	347	23	60.5	327	9	US-09-908-193-39	Sequence 39, Appl
275	23	60.5	113	10	US-09-923-304-7	Sequence 7, Appl	348	23	60.5	328	10	US-09-765-272-8	Sequence 8, Appl
276	23	60.5	115	10	US-09-263-959-352	Sequence 352, App	349	23	60.5	331	10	US-09-815-242-10188	Sequence 10188, A
277	23	60.5	120	9	US-10-103-511-9	Sequence 9, Appl	350	23	60.5	334	10	US-09-815-242-13981	Sequence 13981, A
278	23	60.5	120	10	US-09-805-204-9	Sequence 9, Appl	351	23	60.5	340	9	US-09-908-193-38	Sequence 38, Appl
279	23	60.5	122	10	US-09-925-297-600	Sequence 600, App	352	23	60.5	341	10	US-09-925-302-448	Sequence 448, App
280	23	60.5	131	12	US-10-062-254-220	Sequence 220, App	353	23	60.5	344	10	US-09-815-242-11736	Sequence 11736, A
281	23	60.5	132	10	US-09-764-855-135	Sequence 135, App	354	23	60.5	348	10	US-09-764-870-502	Sequence 502, App
282	23	60.5	133	10	US-09-738-769A-4	Sequence 4, Appl	355	23	60.5	348	12	US-10-027-450-19	Sequence 19, Appl
283	23	60.5	133	10	US-09-867-550-598	Sequence 598, App	356	23	60.5	349	10	US-09-815-242-5073	Sequence 5073, App
284	23	60.5	140	10	US-09-765-272-100	Sequence 100, App	357	23	60.5	350	10	US-09-908-8058-25	Sequence 25, Appl
285	23	60.5	144	9	US-09-854-133-73	Sequence 73, Appl	358	23	60.5	357	10	US-09-791-961-4	Sequence 4, Appl
286	23	60.5	144	10	US-09-738-973-73	Sequence 73, Appl	359	23	60.5	360	10	US-09-893-348-20	Sequence 20, Appl
287	23	60.5	146	10	US-09-923-304-9	Sequence 9, Appl	360	23	60.5	365	9	US-10-028-180-10	Sequence 10, Appl
288	23	60.5	158	10	US-09-734-569-76	Sequence 76, Appl	361	23	60.5	372	9	US-09-789-386-6	Sequence 6, Appl
289	23	60.5	159	10	US-09-881-752A-82	Sequence 82, Appl	362	23	60.5	373	10	US-09-765-205-6	Sequence 24, Appl
290	23	60.5	164	9	US-09-903-327A-10	Sequence 10, Appl	363	23	60.5	373	10	US-09-893-348-24	Sequence 24, Appl
291	23	60.5	166	9	US-09-748-592-2	Sequence 2, Appl	364	23	60.5	376	10	US-09-815-242-11031	Sequence 14031, A
292	23	60.5	167	9	US-09-764-868-775	Sequence 775, App	365	23	60.5	378	9	US-09-971-536-46	Sequence 46, Appl
293	23	60.5	169	10	US-09-903-814-4	Sequence 4, Appl	366	23	60.5	385	9	US-10-028-180-42	Sequence 42, Appl
294	23	60.5	179	10	US-09-834-765-767	Sequence 767, App	367	23	60.5	385	9	US-10-062-254-246	Sequence 246, App
295	23	60.5	190	10	US-09-877-964-2	Sequence 2, Appl	368	23	60.5	389	12	US-09-815-242-12004	Sequence 12004, A
296	23	60.5	190	10	US-09-877-964-4	Sequence 4, Appl	369	23	60.5	407	10	US-09-764-870-375	Sequence 375, App
297	23	60.5	195	10	US-09-005-243-44	Sequence 44, Appl	370	23	60.5	407	10	US-09-815-242-11965	Sequence 11965, A
298	23	60.5	195	10	US-09-224-683-44	Sequence 44, Appl	371	23	60.5	412	9	US-10-029-180-84	Sequence 84, Appl
299	23	60.5	199	10	US-09-893-348-21	Sequence 21, Appl	372	23	60.5	418	10	US-09-796-089-7	Sequence 7, Appl
300	23	60.5	199	10	US-09-893-348-25	Sequence 25, Appl	373	23	60.5	426	9	US-09-464-099A-63	Sequence 63, Appl
301	23	60.5	205	10	US-09-841-132-494	Sequence 494, App	374	23	60.5	431	10	US-09-861-696-63	Sequence 11033, A
302	23	60.5	208	10	US-09-005-243-46	Sequence 46, Appl	375	23	60.5	436	10	US-09-815-242-11033	Sequence 11033, A
303	23	60.5	208	10	US-09-224-683-46	Sequence 46, Appl	376	23	60.5	433	12	US-10-078-929-172	Sequence 172, App
304	23	60.5	220	10	US-09-815-242-10461	Sequence 10461, A	377	23	60.5	433	9	US-09-712-363-258	Sequence 258, App
305	23	60.5	244	10	US-09-815-242-10883	Sequence 10883, A	378	23	60.5	436	9	US-09-764-868-826	Sequence 826, App
306	23	60.5	245	10	US-09-005-243-63	Sequence 63, Appl	379	23	60.5	440	10	US-09-730-525-24	Sequence 24, Appl
307	23	60.5	245	10	US-09-224-683-63	Sequence 63, Appl	380	23	60.5	440	10	US-09-730-917-24	Sequence 24, Appl
308	23	60.5	259	10	US-09-925-299-942	Sequence 942, App	381	23	60.5	440	10	US-09-815-242-11781	Sequence 11781, A
309	23	60.5	266	10	US-09-005-243-57	Sequence 57, Appl	382	23	60.5	455	10	US-09-815-242-10171	Sequence 10171, A
310	23	60.5	266	10	US-09-224-683-57	Sequence 57, Appl	383	23	60.5	470	10	US-09-815-242-10171	Sequence 10171, A
311	23	60.5	269	10	US-09-978-486-3	Sequence 3, Appl	384	23	60.5	471	10	US-09-815-242-13956	Sequence 13956, A

385	23	60.5	481	10	US-09-730-525-6	Sequence 6, Appl	458	23	60.5	719	9	US-09-766-253-7	Sequence 7, Appl
386	23	60.5	481	10	US-09-730-917-6	Sequence 6, Appl	459	23	60.5	727	10	US-09-445-023A-1	Sequence 1, Appl
387	23	60.5	485	10	US-09-815-242-11090	Sequence 11090, A	460	23	60.5	727	10	US-09-445-023A-12	Sequence 12, Appl
388	23	60.5	488	10	US-09-815-242-13245	Sequence 13245, A	461	23	60.5	731	9	US-09-361-630-3	Sequence 3, Appl
389	23	60.5	490	10	US-09-841-132-434	Sequence 434, App	462	23	60.5	737	10	US-09-771-161A-195	Sequence 15, App
390	23	60.5	497	10	US-09-815-242-10378	Sequence 10378, A	463	23	60.5	740	10	US-09-939-408A-15	Sequence 59, Appl
391	23	60.5	498	10	US-09-892-985-12	Sequence 12, Appl	464	23	60.5	767	10	US-09-919-497-55	Sequence 16, Appl
392	23	60.5	500	12	US-10-033-109-4	Sequence 4, Appl	465	23	60.5	781	10	US-09-815-242-11269	Sequence 56, Appl
393	23	60.5	501	9	US-09-918-543-25	Sequence 25, Appl	466	23	60.5	790	10	US-09-866-582-16	Sequence 11269, A
394	23	60.5	501	9	US-09-918-543-27	Sequence 27, Appl	467	23	60.5	792	10	US-09-925-301-1313	Sequence 1313, Ap
395	23	60.5	501	9	US-09-918-543-28	Sequence 28, Appl	468	23	60.5	796	10	US-09-765-772-56	Sequence 56, Appl
396	23	60.5	501	10	US-09-971-611-4	Sequence 4, Appl	469	23	60.5	806	12	US-10-003-405-2	Sequence 2, Appl
397	23	60.5	513	10	US-09-978-486-2	Sequence 14066, A	470	23	60.5	812	10	US-09-815-242-11294	Sequence 11294, A
398	23	60.5	514	10	US-09-925-300-1667	Sequence 2, Appl	471	23	60.5	814	12	US-10-014-501-4	Sequence 4, Appl
399	23	60.5	521	10	US-09-925-300-1667	Sequence 1667, Ap	472	23	60.5	816	10	US-09-815-242-13947	Sequence 13947, A
400	23	60.5	524	10	US-09-912-176-1	Sequence 1, Appl	473	23	60.5	819	9	US-09-766-720B-4	Sequence 4, Appl
401	23	60.5	529	10	US-09-795-693-31	Sequence 31, Appl	474	23	60.5	824	10	US-09-866-582-34	Sequence 34, Appl
402	23	60.5	536	10	US-09-816-028A-19	Sequence 19, Appl	475	23	60.5	832	10	US-09-834-765-2	Sequence 2, Appl
403	23	60.5	536	10	US-09-816-028A-25	Sequence 25, Appl	476	23	60.5	837	10	US-09-815-242-13471	Sequence 13471, A
404	23	60.5	537	9	US-09-486-734A-8	Sequence 8, Appl	477	23	60.5	844	10	US-09-815-242-11497	Sequence 11497, A
405	23	60.5	543	10	US-09-741-659-439	Sequence 439, App	478	23	60.5	844	10	US-09-815-242-11497	Sequence 11497, A
406	23	60.5	548	10	US-09-964-469-4	Sequence 4, Appl	479	23	60.5	855	12	US-10-014-501-2	Sequence 2, Appl
407	23	60.5	548	10	US-09-887-586A-2	Sequence 2, Appl	480	23	60.5	856	10	US-09-815-242-11489	Sequence 11489, A
408	23	60.5	548	10	US-09-887-586A-4	Sequence 4, Appl	481	23	60.5	858	10	US-09-815-242-11396	Sequence 11396, A
409	23	60.5	548	10	US-09-887-586A-6	Sequence 6, Appl	482	23	60.5	893	10	US-09-916-790-5	Sequence 5, Appl
410	23	60.5	548	10	US-09-887-586A-8	Sequence 8, Appl	483	23	60.5	894	12	US-10-060-332-4	Sequence 4, Appl
411	23	60.5	548	10	US-09-887-586A-10	Sequence 10, Appl	484	23	60.5	921	10	US-09-804-474A-4	Sequence 2, Appl
412	23	60.5	548	10	US-09-887-586A-12	Sequence 12, Appl	485	23	60.5	921	12	US-10-054-680-2	Sequence 2, Appl
413	23	60.5	548	10	US-09-903-012-2	Sequence 2, Appl	486	23	60.5	921	12	US-10-054-680-2	Sequence 2, Appl
414	23	60.5	548	10	US-09-903-012-6	Sequence 6, Appl	487	23	60.5	927	10	US-09-804-474A-4	Sequence 4, Appl
415	23	60.5	548	10	US-09-903-012-8	Sequence 8, Appl	488	23	60.5	928	9	US-10-108-605-261	Sequence 2, Appl
416	23	60.5	548	10	US-09-903-012-10	Sequence 10, Appl	489	23	60.5	928	10	US-09-801-574-44	Sequence 44, Appl
417	23	60.5	548	10	US-09-903-012-12	Sequence 12, Appl	490	23	60.5	938	10	US-09-815-242-11016	Sequence 10016, A
418	23	60.5	548	10	US-09-903-012-12	Sequence 12, Appl	491	23	60.5	950	10	US-09-821-987B-4	Sequence 4, Appl
419	23	60.5	548	10	US-09-903-012-12	Sequence 12, Appl	492	23	60.5	953	10	US-09-728-721-43	Sequence 21, Appl
420	23	60.5	553	10	US-09-964-469-2	Sequence 2, Appl	493	23	60.5	970	12	US-10-105-925-2	Sequence 43, Appl
421	23	60.5	553	12	US-10-044-205A-2	Sequence 2, Appl	494	23	60.5	1036	10	US-09-771-161A-555	Sequence 255, App
422	23	60.5	553	12	US-09-815-242-11735	Sequence 11735, A	495	23	60.5	1036	10	US-09-771-161A-555	Sequence 255, App
423	23	60.5	555	10	US-09-815-242-13709	Sequence 13709, A	496	23	60.5	1036	10	US-09-826-660-21	Sequence 21, Appl
424	23	60.5	560	10	US-09-820-952A-3	Sequence 3, Appl	497	23	60.5	1163	10	US-09-893-348-18	Sequence 18, Appl
425	23	60.5	560	10	US-09-820-952A-3	Sequence 3, Appl	498	23	60.5	1163	10	US-09-789-386-2	Sequence 2, Appl
426	23	60.5	567	10	US-09-820-953-3	Sequence 3, Appl	499	23	60.5	1163	10	US-09-893-348-23	Sequence 23, Appl
427	23	60.5	567	10	US-09-820-953-3	Sequence 3, Appl	500	23	60.5	1163	10	US-09-893-348-23	Sequence 23, Appl
428	23	60.5	578	10	US-09-815-242-13563	Sequence 13563, A	501	23	60.5	1163	10	US-09-893-348-23	Sequence 23, Appl
429	23	60.5	583	12	US-10-021-657-2	Sequence 12474, A	502	23	60.5	1163	10	US-09-893-348-23	Sequence 23, Appl
430	23	60.5	592	10	US-09-905-173-30	Sequence 30, Appl	503	23	60.5	1163	10	US-09-893-348-23	Sequence 23, Appl
431	23	60.5	608	10	US-09-803-589-2	Sequence 8, Appl	504	23	60.5	1163	10	US-09-893-348-23	Sequence 23, Appl
432	23	60.5	608	10	US-09-803-589-2	Sequence 8, Appl	505	23	60.5	1163	10	US-09-893-348-23	Sequence 23, Appl
433	23	60.5	608	10	US-09-803-589-2	Sequence 8, Appl	506	23	60.5	1163	10	US-09-893-348-23	Sequence 23, Appl
434	23	60.5	608	10	US-09-803-589-2	Sequence 8, Appl	507	23	60.5	1163	10	US-09-893-348-23	Sequence 23, Appl
435	23	60.5	613	9	US-09-903-327A-14	Sequence 14, Appl	508	23	60.5	1163	10	US-09-893-348-23	Sequence 23, Appl
436	23	60.5	613	10	US-09-815-242-10225	Sequence 10225, A	509	23	60.5	1163	10	US-09-893-348-23	Sequence 23, Appl
437	23	60.5	613	10	US-09-815-242-10225	Sequence 10225, A	510	23	60.5	1163	10	US-09-893-348-23	Sequence 23, Appl
438	23	60.5	620	12	US-10-054-680-4	Sequence 12886, A	511	23	60.5	1163	10	US-09-893-348-23	Sequence 23, Appl
439	23	60.5	624	10	US-09-834-765-5	Sequence 5, Appl	512	23	60.5	1163	10	US-09-893-348-23	Sequence 23, Appl
440	23	60.5	625	10	US-09-834-765-5	Sequence 5, Appl	513	23	60.5	1163	10	US-09-893-348-23	Sequence 23, Appl
441	23	60.5	629	10	US-09-741-669-473	Sequence 762, App	514	23	60.5	1163	10	US-09-893-348-23	Sequence 23, Appl
442	23	60.5	629	10	US-09-815-242-10405	Sequence 10405, A	515	23	60.5	1163	10	US-09-893-348-23	Sequence 23, Appl
443	23	60.5	632	9	US-09-971-556-42	Sequence 11063, A	516	23	60.5	1163	10	US-09-893-348-23	Sequence 23, Appl
444	23	60.5	632	9	US-09-971-556-42	Sequence 11063, A	517	23	60.5	1163	10	US-09-893-348-23	Sequence 23, Appl
445	23	60.5	637	10	US-09-815-242-11225	Sequence 11225, A	518	23	60.5	1163	10	US-09-893-348-23	Sequence 23, Appl
446	23	60.5	637	10	US-09-815-242-11225	Sequence 11225, A	519	23	60.5	1163	10	US-09-893-348-23	Sequence 23, Appl
447	23	60.5	637	10	US-09-815-242-11225	Sequence 11225, A	520	23	60.5	1163	10	US-09-893-348-23	Sequence 23, Appl
448	23	60.5	640	9	US-09-983-204-16	Sequence 16, Appl	521	23	60.5	1163	10	US-09-893-348-23	Sequence 23, Appl
449	23	60.5	640	9	US-10-133-157-5	Sequence 5, Appl	522	23	60.5	1163	10	US-09-893-348-23	Sequence 23, Appl
450	23	60.5	649	10	US-09-815-242-4979	Sequence 4979, Ap	523	23	60.5	1163	10	US-09-893-348-23	Sequence 23, Appl
451	23	60.5	649	10	US-09-815-242-10952	Sequence 10952, A	524	23	60.5	1163	10	US-09-893-348-23	Sequence 23, Appl
452	23	60.5	678	9	US-09-712-363-158	Sequence 158, App	525	23	60.5	1163	10	US-09-893-348-23	Sequence 23, Appl
453	23	60.5	684	9	US-09-925-302-555	Sequence 555, App	526	23	60.5	1163	10	US-09-893-348-23	Sequence 23, Appl
454	23	60.5	699	9	US-09-843-676-52	Sequence 52, Appl	527	23	60.5	1163	10	US-09-893-348-23	Sequence 23, Appl
455	23	60.5	699	9	US-09-766-253-52	Sequence 52, Appl	528	23	60.5	1163	10	US-09-893-348-23	Sequence 23, Appl
456	23	60.5	701	10	US-09-815-242-13411	Sequence 13411, A	529	23	60.5	1163	10	US-09-893-348-23	Sequence 23, Appl
457	23	60.5	719	9	US-09-843-676-7	Sequence 7, Appl	530	23	60.5	1163	10	US-09-893-348-23	Sequence 23, Appl

531	22	57.9	47	10	US-09-864-761-47008	Sequence 47008, A	604	22	57.9	271	10	US-09-815-242-14072	Sequence 14072, A
532	22	57.9	48	10	US-09-864-761-48373	Sequence 48373, A	605	22	57.9	288	10	US-09-925-297-617	Sequence 617, App
533	22	57.9	49	10	US-09-864-761-43229	Sequence 43229, A	606	22	57.9	289	10	US-09-925-301-1408	Sequence 1408, App
534	22	57.9	49	10	US-09-764-860-526	Sequence 526, App	607	22	57.9	289	10	US-09-815-242-4948	Sequence 4948, App
535	22	57.9	53	10	US-09-864-761-40970	Sequence 40970, A	608	22	57.9	293	10	US-09-815-242-10488	Sequence 10488, A
536	22	57.9	53	10	US-09-864-761-44080	Sequence 44080, A	609	22	57.9	295	10	US-09-547-267-1	Sequence 1, Appl
537	22	57.9	55	10	US-09-864-761-40105	Sequence 40105, A	610	22	57.9	299	10	US-09-861-451A-36	Sequence 36, Appl
538	22	57.9	55	10	US-09-864-761-44639	Sequence 44639, A	611	22	57.9	300	10	US-09-815-242-5185	Sequence 5185, App
539	22	57.9	65	10	US-09-815-242-12049	Sequence 12049, A	612	22	57.9	301	10	US-09-925-301-1354	Sequence 1354, App
540	22	57.9	66	10	US-09-864-761-48591	Sequence 48591, A	613	22	57.9	301	10	US-09-925-300-1021	Sequence 1021, App
541	22	57.9	72	10	US-09-864-761-46760	Sequence 46760, A	614	22	57.9	303	10	US-09-547-267-3	Sequence 3, Appl
542	22	57.9	73	10	US-09-881-752A-240	Sequence 240, App	615	22	57.9	307	12	US-10-052-586-262	Sequence 262, App
543	22	57.9	80	12	US-10-052-586-576	Sequence 576, App	616	22	57.9	313	10	US-09-815-242-11097	Sequence 12097, A
544	22	57.9	81	10	US-09-764-877-1340	Sequence 1340, App	617	22	57.9	313	10	US-09-801-368-58	Sequence 58, Appl
545	22	57.9	89	10	US-09-764-877-1676	Sequence 1676, App	618	22	57.9	318	10	US-09-765-205-36	Sequence 36, Appl
546	22	57.9	104	10	US-09-815-242-5168	Sequence 5168, App	619	22	57.9	319	10	US-09-788-626-6	Sequence 6, Appl
547	22	57.9	105	10	US-09-107-058-8	Sequence 8, Appl	620	22	57.9	320	10	US-09-965-703-15	Sequence 15, Appl
548	22	57.9	105	10	US-09-761-117-8	Sequence 8, Appl	621	22	57.9	321	10	US-09-815-242-12150	Sequence 12150, A
549	22	57.9	107	10	US-09-864-761-48250	Sequence 48250, A	622	22	57.9	321	10	US-09-815-242-12776	Sequence 12776, A
550	22	57.9	116	10	US-09-867-550-146	Sequence 146, App	623	22	57.9	321	10	US-09-815-242-13064	Sequence 13064, A
551	22	57.9	123	9	US-10-112-793-19	Sequence 19, Appl	624	22	57.9	323	10	US-09-815-242-12514	Sequence 12514, A
552	22	57.9	128	12	US-10-078-929-4	Sequence 4, Appl	625	22	57.9	323	10	US-09-815-242-12857	Sequence 12857, A
553	22	57.9	130	10	US-09-791-171-64	Sequence 64, Appl	626	22	57.9	325	8	US-08-829-402-2	Sequence 2, Appl
554	22	57.9	144	10	US-09-764-847-923	Sequence 923, App	627	22	57.9	327	10	US-09-815-242-5720	Sequence 5720, App
555	22	57.9	146	10	US-09-815-242-4968	Sequence 4968, App	628	22	57.9	329	10	US-09-815-242-11142	Sequence 11142, A
556	22	57.9	149	10	US-09-815-242-10816	Sequence 10816, A	629	22	57.9	332	10	US-09-962-618-2	Sequence 2, Appl
557	22	57.9	150	10	US-09-864-761-39141	Sequence 39141, A	630	22	57.9	333	10	US-09-815-242-12481	Sequence 12481, A
558	22	57.9	152	10	US-09-798-029-40	Sequence 40, Appl	631	22	57.9	334	9	US-09-971-536-57	Sequence 57, Appl
559	22	57.9	152	10	US-09-799-777-17	Sequence 17, Appl	632	22	57.9	334	10	US-09-965-703-13	Sequence 13, Appl
560	22	57.9	156	10	US-09-925-297-680	Sequence 680, App	633	22	57.9	336	10	US-09-864-761-43157	Sequence 43157, A
561	22	57.9	162	10	US-09-925-302-593	Sequence 593, App	634	22	57.9	340	10	US-09-842-628-2	Sequence 2, Appl
562	22	57.9	164	10	US-09-815-242-13509	Sequence 13509, A	635	22	57.9	340	10	US-09-815-242-11743	Sequence 11743, A
563	22	57.9	163	10	US-09-815-242-13630	Sequence 13630, A	636	22	57.9	341	10	US-09-086-618-25	Sequence 25, Appl
564	22	57.9	170	10	US-09-864-761-34145	Sequence 34145, A	637	22	57.9	341	10	US-09-835-684-7	Sequence 7, Appl
565	22	57.9	172	10	US-09-902-331-10	Sequence 10, Appl	638	22	57.9	341	10	US-09-880-371-7	Sequence 11, Appl
566	22	57.9	174	10	US-09-989-903-16	Sequence 16, Appl	639	22	57.9	341	10	US-09-879-248-11	Sequence 11, Appl
567	22	57.9	176	10	US-09-902-331-6	Sequence 6, Appl	640	22	57.9	342	10	US-09-770-693-5	Sequence 5, Appl
568	22	57.9	181	10	US-09-811-284-156	Sequence 156, App	641	22	57.9	342	9	US-09-995-225-14	Sequence 14, App
569	22	57.9	181	10	US-09-841-132-295	Sequence 295, App	642	22	57.9	343	10	US-09-925-300-1343	Sequence 1343, App
570	22	57.9	183	10	US-09-815-242-11840	Sequence 11840, A	643	22	57.9	343	10	US-09-815-242-5065	Sequence 5065, App
571	22	57.9	183	12	US-10-033-109-2	Sequence 2, Appl	644	22	57.9	344	10	US-09-925-297-735	Sequence 735, App
572	22	57.9	184	10	US-09-379-931-9	Sequence 9, Appl	645	22	57.9	355	9	US-09-947-953-2	Sequence 2, Appl
573	22	57.9	187	10	US-09-917-265-102	Sequence 102, App	646	22	57.9	362	10	US-09-946-785-2	Sequence 2, Appl
574	22	57.9	191	10	US-09-841-132-391	Sequence 391, App	647	22	57.9	363	10	US-09-922-261-81	Sequence 81, Appl
575	22	57.9	196	10	US-09-764-869-1004	Sequence 1004, App	648	22	57.9	365	10	US-09-815-242-5441	Sequence 5441, App
576	22	57.9	197	10	US-09-917-265-36	Sequence 36, Appl	649	22	57.9	370	9	US-10-117-846-8	Sequence 8, Appl
577	22	57.9	197	10	US-09-917-265-50	Sequence 50, Appl	650	22	57.9	370	10	US-09-833-790-253	Sequence 253, App
578	22	57.9	197	12	US-10-079-616-4	Sequence 4, Appl	651	22	57.9	389	10	US-09-925-301-1256	Sequence 1256, App
579	22	57.9	200	10	US-09-870-756-38	Sequence 38, Appl	652	22	57.9	389	10	US-09-731-872-414	Sequence 414, App
580	22	57.9	200	10	US-09-874-585B-38	Sequence 38, Appl	653	22	57.9	396	9	US-10-155-613-2	Sequence 2, Appl
581	22	57.9	206	12	US-10-078-929-6	Sequence 6, Appl	654	22	57.9	396	10	US-09-864-761-43232	Sequence 43232, A
582	22	57.9	208	12	US-10-078-929-6	Sequence 6, Appl	655	22	57.9	397	10	US-09-842-987-2	Sequence 2, Appl
583	22	57.9	221	10	US-09-925-301-946	Sequence 946, App	656	22	57.9	397	10	US-09-815-242-12144	Sequence 12144, A
584	22	57.9	221	10	US-09-811-284-158	Sequence 158, App	657	22	57.9	397	10	US-09-925-637-58	Sequence 58, App
585	22	57.9	222	10	US-09-917-265-33	Sequence 33, Appl	658	22	57.9	401	10	US-09-925-300-1558	Sequence 1550, App
586	22	57.9	222	10	US-09-917-265-47	Sequence 47, Appl	659	22	57.9	402	10	US-09-799-777-71	Sequence 71, Appl
587	22	57.9	222	10	US-09-917-265-105	Sequence 105, App	660	22	57.9	407	10	US-09-925-301-1057	Sequence 1057, App
588	22	57.9	222	12	US-10-079-616-3	Sequence 3, Appl	661	22	57.9	410	9	US-10-011-445-55	Sequence 55, Appl
589	22	57.9	223	10	US-09-815-242-13328	Sequence 13328, A	662	22	57.9	412	10	US-09-922-261-77	Sequence 77, Appl
590	22	57.9	232	10	US-09-841-132-302	Sequence 302, App	663	22	57.9	412	10	US-09-965-703-11	Sequence 11, Appl
591	22	57.9	232	10	US-09-841-132-392	Sequence 392, App	664	22	57.9	412	10	US-09-965-703-12	Sequence 12, Appl
592	22	57.9	232	10	US-09-841-132-92	Sequence 448, App	665	22	57.9	419	10	US-09-750-964-6	Sequence 6, Appl
593	22	57.9	238	9	US-09-727-855B-9	Sequence 9, Appl	666	22	57.9	421	10	US-09-815-242-13330	Sequence 13330, A
594	22	57.9	238	10	US-09-452-239-28	Sequence 28, Appl	667	22	57.9	421	10	US-09-793-306-163	Sequence 163, App
595	22	57.9	238	12	US-10-007-693-124	Sequence 124, App	668	22	57.9	424	10	US-09-925-302-583	Sequence 583, App
596	22	57.9	248	10	US-09-815-242-12098	Sequence 12098, A	669	22	57.9	426	9	US-10-004-381-11	Sequence 41, Appl
597	22	57.9	254	9	US-09-981-353-70	Sequence 70, Appl	670	22	57.9	426	10	US-09-815-242-110040	Sequence 10040, A
598	22	57.9	260	10	US-09-925-299-919	Sequence 919, App	671	22	57.9	426	10	US-09-815-242-13306	Sequence 13306, A
599	22	57.9	261	10	US-09-864-761-34748	Sequence 34748, A	672	22	57.9	427	10	US-09-764-871-482	Sequence 482, App
600	22	57.9	261	12	US-10-062-254-140	Sequence 140, App	673	22	57.9	427	10	US-09-835-881-6	Sequence 6, Appl
601	22	57.9	261	12	US-10-062-254-142	Sequence 142, App	674	22	57.9	428	10	US-09-835-881-6	Sequence 6, Appl
602	22	57.9	267	10	US-09-815-242-5328	Sequence 5328, App	675	22	57.9	429	12	US-10-042-311-6	Sequence 12, Appl
603	22	57.9					676	22	57.9				

677	22	57.9	429	10	US-09-866-582-30	Sequence 30, Appl	750	22	57.9	533	10	US-09-917-265-39	Sequence 39, Appl
678	22	57.9	432	10	US-09-764-864-1184	Sequence 1184, Ap	751	22	57.9	533	10	US-09-917-265-62	Sequence 62, Appl
679	22	57.9	440	10	US-09-965-703-71	Sequence 71, Appl	752	22	57.9	535	10	US-09-006-298-23	Sequence 23, Appl
680	22	57.9	441	10	US-09-833-017-4	Sequence 4, Appl	753	22	57.9	538	10	US-09-758-664-2	Sequence 2, Appl
681	22	57.9	442	10	US-09-815-242-13237	Sequence 13237, A	754	22	57.9	538	10	US-09-965-313-12	Sequence 115, Appl
682	22	57.9	442	10	US-09-815-242-13702	Sequence 13702, A	755	22	57.9	538	10	US-09-923-466-115	Sequence 2, Appl
683	22	57.9	446	10	US-09-815-242-4872	Sequence 4872, Ap	756	22	57.9	538	10	US-09-825-61A-2	Sequence 10, Appl
684	22	57.9	453	10	US-09-784-810A-12	Sequence 12, Appl	757	22	57.9	539	10	US-09-945-825-10	Sequence 10, Appl
685	22	57.9	458	9	US-09-843-905A-15	Sequence 15, Appl	758	22	57.9	546	10	US-09-901-884-5	Sequence 5, Appl
686	22	57.9	461	10	US-09-815-242-10810	Sequence 10810, A	759	22	57.9	551	10	US-09-897-214-8	Sequence 8, Appl
687	22	57.9	462	10	US-09-977-727-3	Sequence 3, Appl	760	22	57.9	552	12	US-10-029-654-8	Sequence 8, Appl
688	22	57.9	465	10	US-09-925-302-771	Sequence 771, App	761	22	57.9	553	10	US-09-815-242-5497	Sequence 5497, Ap
689	22	57.9	474	9	US-09-736-457-1812	Sequence 1812, Ap	762	22	57.9	553	10	US-09-815-242-12429	Sequence 12429, A
690	22	57.9	474	9	US-09-902-941-1812	Sequence 1812, Ap	763	22	57.9	558	9	US-09-850-991-2	Sequence 2, Appl
691	22	57.9	475	12	US-10-142-373-2	Sequence 2, Appl	764	22	57.9	559	12	US-10-001-851-20	Sequence 20, Appl
692	22	57.9	478	10	US-09-765-111A-27	Sequence 27, Appl	765	22	57.9	559	12	US-10-001-851-21	Sequence 21, Appl
693	22	57.9	479	9	US-10-004-981-38	Sequence 38, Appl	766	22	57.9	559	12	US-10-001-851-22	Sequence 22, Appl
694	22	57.9	479	10	US-09-815-242-10759	Sequence 10759, A	767	22	57.9	559	12	US-10-001-851-23	Sequence 23, Appl
695	22	57.9	481	1	US-08-976-063C-28	Sequence 28, Appl	768	22	57.9	559	12	US-10-001-851-24	Sequence 24, Appl
696	22	57.9	481	10	US-09-815-242-10158	Sequence 10158, A	769	22	57.9	561	10	US-09-925-301-1006	Sequence 1006, Ap
697	22	57.9	486	9	US-09-764-868-1149	Sequence 1149, Ap	770	22	57.9	565	10	US-09-768-826-58	Sequence 58, Appl
698	22	57.9	486	9	US-10-005-057A-13	Sequence 13, Appl	771	22	57.9	568	10	US-09-758-664-4	Sequence 4, Appl
699	22	57.9	489	9	US-09-764-868-776	Sequence 776, App	772	22	57.9	578	12	US-09-825-561A-65	Sequence 65, Appl
700	22	57.9	493	10	US-09-815-242-12052	Sequence 12052, A	773	22	57.9	580	10	US-10-074-527-8	Sequence 8, Appl
701	22	57.9	494	12	US-10-033-109-6	Sequence 6, Appl	774	22	57.9	590	10	US-09-943-671-25	Sequence 25, Appl
702	22	57.9	495	10	US-09-923-922-2	Sequence 2, Appl	775	22	57.9	595	10	US-09-826-212-9	Sequence 9, Appl
703	22	57.9	496	10	US-09-924-417-65	Sequence 65, Appl	776	22	57.9	595	10	US-09-921-667-6	Sequence 6, Appl
704	22	57.9	501	8	US-08-635-967-2	Sequence 2, Appl	777	22	57.9	595	10	US-09-921-667-6	Sequence 6, Appl
705	22	57.9	503	10	US-09-990-337-3	Sequence 3, Appl	778	22	57.9	596	10	US-09-935-727-11	Sequence 11, Appl
706	22	57.9	505	9	US-09-712-363-275	Sequence 275, Appl	779	22	57.9	597	10	US-09-815-242-5244	Sequence 5244, Ap
707	22	57.9	505	10	US-09-738-878-4	Sequence 4, Appl	780	22	57.9	600	10	US-09-815-242-11455	Sequence 11455, A
708	22	57.9	505	10	US-09-738-878-5	Sequence 5, Appl	781	22	57.9	601	10	US-09-801-368-158	Sequence 158, Appl
709	22	57.9	505	10	US-09-738-878-6	Sequence 6, Appl	782	22	57.9	601	10	US-09-887-586A-42	Sequence 42, Appl
710	22	57.9	505	10	US-09-738-878-7	Sequence 7, Appl	783	22	57.9	604	10	US-09-903-012-42	Sequence 42, Appl
711	22	57.9	505	10	US-09-738-878-8	Sequence 8, Appl	784	22	57.9	606	10	US-09-815-242-12525	Sequence 12525, A
712	22	57.9	505	10	US-09-765-111A-16	Sequence 16, Appl	785	22	57.9	606	10	US-09-825-561A-25	Sequence 25, Appl
713	22	57.9	505	12	US-10-163-381-4	Sequence 4, Appl	786	22	57.9	612	12	US-10-001-851-35	Sequence 35, Appl
714	22	57.9	505	12	US-10-163-381-5	Sequence 5, Appl	787	22	57.9	615	9	US-09-908-193-27	Sequence 27, Appl
715	22	57.9	505	12	US-10-163-381-6	Sequence 6, Appl	788	22	57.9	615	9	US-09-908-193-27	Sequence 27, Appl
716	22	57.9	505	12	US-10-163-381-7	Sequence 7, Appl	789	22	57.9	619	12	US-10-007-693-77	Sequence 77, Appl
717	22	57.9	505	12	US-10-163-381-8	Sequence 8, Appl	790	22	57.9	620	10	US-09-815-242-11472	Sequence 11472, A
718	22	57.9	506	12	US-10-109-886-6	Sequence 6, Appl	791	22	57.9	623	10	US-09-862-007-2	Sequence 2, Appl
719	22	57.9	510	10	US-09-923-246-89	Sequence 89, Appl	792	22	57.9	623	10	US-09-862-007-2	Sequence 2, Appl
720	22	57.9	511	10	US-09-912-020-259	Sequence 259, Appl	793	22	57.9	632	10	US-09-925-300-1516	Sequence 1516, Ap
721	22	57.9	511	10	US-09-917-265-44	Sequence 44, Appl	794	22	57.9	636	9	US-09-815-242-10912	Sequence 10912, A
722	22	57.9	511	10	US-09-917-265-67	Sequence 67, Appl	795	22	57.9	637	10	US-10-000-776-12	Sequence 12, Appl
723	22	57.9	513	10	US-09-949-780-2	Sequence 2, Appl	796	22	57.9	637	10	US-09-815-242-13866	Sequence 13866, A
724	22	57.9	513	10	US-09-833-745-58	Sequence 58, Appl	797	22	57.9	638	9	US-10-072-621-10	Sequence 10, Appl
725	22	57.9	513	10	US-09-965-703-59	Sequence 59, Appl	798	22	57.9	639	10	US-09-789-417-2	Sequence 2, Appl
726	22	57.9	518	10	US-09-746-390-4	Sequence 4, Appl	799	22	57.9	639	10	US-09-792-451-2	Sequence 15, Appl
727	22	57.9	519	10	US-09-923-246-85	Sequence 85, Appl	800	22	57.9	639	10	US-09-835-996A-15	Sequence 15, Appl
728	22	57.9	522	10	US-09-965-313-7	Sequence 7, Appl	801	22	57.9	639	12	US-10-052-586-364	Sequence 364, App
729	22	57.9	531	9	US-10-047-260-32	Sequence 32, Appl	802	22	57.9	646	9	US-10-072-621-9	Sequence 9, Appl
730	22	57.9	532	9	US-09-992-598-381	Sequence 381, App	803	22	57.9	646	9	US-10-154-452-4	Sequence 4, Appl
731	22	57.9	532	9	US-09-989-293A-381	Sequence 381, App	804	22	57.9	646	9	US-10-154-452-8	Sequence 8, Appl
732	22	57.9	532	9	US-10-063-547-72	Sequence 72, Appl	805	22	57.9	646	9	US-09-833-555-20	Sequence 20, Appl
733	22	57.9	532	9	US-09-989-735-381	Sequence 381, App	806	22	57.9	660	10	US-09-853-033-6	Sequence 6, Appl
734	22	57.9	532	9	US-09-990-444-381	Sequence 381, App	807	22	57.9	660	10	US-09-853-033-8	Sequence 8, Appl
735	22	57.9	532	10	US-09-989-723-381	Sequence 381, App	808	22	57.9	681	9	US-09-969-384-25	Sequence 25, Appl
736	22	57.9	532	10	US-09-989-723-381	Sequence 381, App	809	22	57.9	684	10	US-09-815-242-10712	Sequence 10712, A
737	22	57.9	532	10	US-09-988-279-381	Sequence 381, App	810	22	57.9	688	9	US-09-931-007A-1	Sequence 1, Appl
738	22	57.9	532	10	US-09-989-727-381	Sequence 381, App	811	22	57.9	688	10	US-09-815-242-11416	Sequence 11416, A
739	22	57.9	532	10	US-09-989-731-381	Sequence 381, App	812	22	57.9	688	10	US-09-815-242-11576	Sequence 11576, A
740	22	57.9	532	10	US-09-989-732-381	Sequence 381, App	813	22	57.9	690	10	US-09-934-477-47	Sequence 47, Appl
741	22	57.9	532	10	US-09-991-073-381	Sequence 381, App	814	22	57.9	691	9	US-09-935-868-20	Sequence 20, Appl
742	22	57.9	532	10	US-09-990-442-381	Sequence 381, App	815	22	57.9	693	9	US-10-029-217A-4	Sequence 4, Appl
743	22	57.9	532	10	US-09-991-163-381	Sequence 381, App	816	22	57.9	694	9	US-09-935-868-18	Sequence 18, Appl
744	22	57.9	532	10	US-09-993-604-381	Sequence 381, App	817	22	57.9	694	9	US-09-935-868-22	Sequence 22, Appl
745	22	57.9	532	10	US-09-990-456-381	Sequence 381, App	818	22	57.9	694	10	US-09-934-477-49	Sequence 49, Appl
746	22	57.9	532	10	US-09-989-721-381	Sequence 381, App	819	22	57.9	708	10	US-09-969-384-13	Sequence 13, Appl
747	22	57.9	532	12	US-10-006-867-72	Sequence 72, Appl	820	22	57.9	711	9	US-09-969-384-13	Sequence 13, Appl
748	22	57.9	532	12	US-10-006-867-72	Sequence 72, Appl	821	22	57.9	711	9	US-10-124-800-18	Sequence 18, Appl
749	22	57.9	533	9	US-09-808-602-87	Sequence 87, Appl	822	22	57.9	711	10	US-09-934-477-35	Sequence 35, Appl



823	22	57.9	714	9	US-09-712-363-149	Sequence 149, App	896	22	57.9	976	10	US-09-801-366-248	Sequence 248, App
824	22	57.9	718	10	US-09-801-368-388	Sequence 388, App	897	22	57.9	980	10	US-09-888-615-65	Sequence 65, App
825	22	57.9	718	10	US-09-764-864-1579	Sequence 1579, Ap	898	22	57.9	1051	10	US-09-911-888-14	Sequence 14, App
826	22	57.9	721	9	US-10-121-032-19	Sequence 19, Appl	899	22	57.9	1074	10	US-09-509-1966-2	Sequence 2, Appl
827	22	57.9	728	9	US-09-712-363-167	Sequence 167, App	900	22	57.9	1115	10	US-09-771-161A-23	Sequence 253, App
828	22	57.9	740	9	US-10-117-846-14	Sequence 14, Appl	901	22	57.9	1137	12	US-10-052-586-518	Sequence 518, App
829	22	57.9	741	9	US-10-117-846-6	Sequence 6, Appl	902	22	57.9	1179	10	US-09-815-242-1362	Sequence 1362, A
830	22	57.9	745	10	US-09-833-790-432	Sequence 432, App	903	22	57.9	1179	10	US-09-815-242-13608	Sequence 13608, A
831	22	57.9	745	10	US-09-919-935-2	Sequence 2, Appl	904	22	57.9	1182	9	US-09-990-046-7	Sequence 7, Appl
832	22	57.9	746	9	US-10-041-007-39	Sequence 39, Appl	905	22	57.9	1184	9	US-09-978-295A-425	Sequence 425, App
833	22	57.9	765	9	US-10-086-135-22	Sequence 22, Appl	906	22	57.9	1184	9	US-09-978-697-425	Sequence 425, App
834	22	57.9	776	9	US-09-935-868-36	Sequence 36, Appl	907	22	57.9	1184	9	US-09-978-192A-425	Sequence 425, App
835	22	57.9	776	9	US-09-935-868-40	Sequence 40, Appl	908	22	57.9	1184	9	US-09-999-832A-425	Sequence 425, App
836	22	57.9	776	9	US-09-935-868-44	Sequence 44, Appl	909	22	57.9	1187	10	US-09-768-436-4	Sequence 4, Appl
837	22	57.9	777	10	US-09-765-111A-2	Sequence 2, Appl	910	22	57.9	1203	9	US-09-990-046-2	Sequence 2, Appl
838	22	57.9	778	9	US-09-935-868-46	Sequence 46, Appl	911	22	57.9	1203	9	US-09-999-280A-2	Sequence 2, Appl
839	22	57.9	778	9	US-09-935-868-50	Sequence 50, Appl	912	22	57.9	1261	10	US-10-147-268-2	Sequence 11963, A
840	22	57.9	780	9	US-09-935-868-34	Sequence 34, Appl	913	22	57.9	1261	12	US-10-147-268-2	Sequence 2, Appl
841	22	57.9	780	9	US-09-935-868-38	Sequence 38, Appl	914	22	57.9	1261	10	US-09-864-761-46846	Sequence 46846, A
842	22	57.9	782	9	US-09-935-868-42	Sequence 42, Appl	915	22	57.9	1284	10	US-09-119-855-11	Sequence 11, Appl
843	22	57.9	782	9	US-09-935-868-48	Sequence 48, Appl	916	22	57.9	1290	10	US-09-881-752A-138	Sequence 138, App
844	22	57.9	782	9	US-09-935-868-52	Sequence 52, Appl	917	22	57.9	1375	10	US-09-740-274-4	Sequence 4, Appl
845	22	57.9	783	10	US-09-864-761-38209	Sequence 38209, A	918	22	57.9	1379	10	US-09-862-179A-44	Sequence 44, Appl
846	22	57.9	784	9	US-09-935-868-30	Sequence 30, Appl	919	22	57.9	1400	10	US-09-764-176-7	Sequence 7, Appl
847	22	57.9	793	9	US-09-935-868-32	Sequence 32, Appl	920	22	57.9	1426	9	US-10-024-632-17	Sequence 17, Appl
848	22	57.9	795	9	US-10-041-007-37	Sequence 37, Appl	921	22	57.9	1464	10	US-09-746-330-2	Sequence 2, Appl
849	22	57.9	803	9	US-10-029-382-2	Sequence 2, Appl	922	22	57.9	1475	10	US-09-740-274-2	Sequence 2, Appl
850	22	57.9	803	10	US-09-391-340-4	Sequence 4, Appl	923	22	57.9	1476	10	US-09-817-514A-4	Sequence 4, Appl
851	22	57.9	803	10	US-09-948-369-4	Sequence 4, Appl	924	22	57.9	1491	10	US-09-815-242-12162	Sequence 5568, Ap
852	22	57.9	803	10	US-09-948-369-14	Sequence 14, Appl	925	22	57.9	1502	10	US-09-815-242-12178	Sequence 12162, A
853	22	57.9	803	10	US-09-948-369-16	Sequence 16, Appl	926	22	57.9	1502	10	US-09-841-132-12178	Sequence 178, App
854	22	57.9	803	12	US-10-034-849-2	Sequence 2, Appl	927	22	57.9	1531	12	US-10-007-653-98	Sequence 322, App
855	22	57.9	803	12	US-10-034-621-2	Sequence 2, Appl	928	22	57.9	1536	10	US-09-801-368-322	Sequence 98, App
856	22	57.9	807	9	US-09-847-046-2	Sequence 2, Appl	929	22	57.9	1602	10	US-09-778-927A-59	Sequence 59, Appl
857	22	57.9	808	9	US-10-055-364-38	Sequence 38, Appl	930	22	57.9	1616	9	US-09-887-540A-2	Sequence 2, Appl
858	22	57.9	811	9	US-09-765-111A-23	Sequence 23, Appl	931	22	57.9	1704	10	US-09-864-761-46862	Sequence 46862, A
859	22	57.9	814	9	US-10-041-007-35	Sequence 35, Appl	932	22	57.9	1812	9	US-09-775-938A-38	Sequence 38, Appl
860	22	57.9	816	10	US-09-815-242-11849	Sequence 11849, A	933	22	57.9	1832	9	US-10-014-717-4	Sequence 4, Appl
861	22	57.9	825	10	US-09-785-934-2	Sequence 2, Appl	934	22	57.9	1863	9	US-09-734-672-2	Sequence 2, Appl
862	22	57.9	825	10	US-09-921-667-16	Sequence 16, Appl	935	22	57.9	1863	9	US-09-734-672-6	Sequence 6, Appl
863	22	57.9	825	10	US-09-904-245-2	Sequence 2, Appl	936	22	57.9	1863	9	US-09-734-672-6	Sequence 6, Appl
864	22	57.9	834	10	US-09-826-752-4	Sequence 4, Appl	937	22	57.9	1909	10	US-09-801-368-390	Sequence 390, App
865	22	57.9	840	10	US-09-765-111A-4	Sequence 4, Appl	938	22	57.9	2190	10	US-09-727-384-6	Sequence 6, Appl
866	22	57.9	842	10	US-09-771-161A-163	Sequence 163, App	939	22	57.9	2308	12	US-10-000-954-2	Sequence 2, Appl
867	22	57.9	845	10	US-09-815-242-4972	Sequence 4972, Ap	940	22	57.9	2436	10	US-09-769-693-8	Sequence 8, Appl
868	22	57.9	845	10	US-09-815-242-10848	Sequence 10848, A	941	22	57.9	2843	8	US-08-681-219-32	Sequence 32, Appl
869	22	57.9	845	10	US-09-949-192-15	Sequence 15, App	942	22	57.9	2843	9	US-09-987-482-1	Sequence 1, Appl
870	22	57.9	848	10	US-09-841-132-192	Sequence 19, App	943	22	57.9	2910	9	US-10-124-800-2	Sequence 2, Appl
871	22	57.9	861	9	US-09-826-312-10	Sequence 10, Appl	944	22	57.9	4019	9	US-09-854-133-425	Sequence 425, App
872	22	57.9	862	9	US-10-041-007-41	Sequence 41, Appl	945	22	57.9	4019	10	US-09-738-973-425	Sequence 425, App
873	22	57.9	862	10	US-09-887-586A-44	Sequence 44, Appl	946	22	57.9	4636	9	US-09-870-759-128	Sequence 128, App
874	22	57.9	862	10	US-09-903-012-44	Sequence 44, Appl	947	22	57.9	4636	9	US-09-976-059-14	Sequence 14, Appl
875	22	57.9	862	10	US-09-815-242-5274	Sequence 5274, Ap	948	22	57.9	4999	9	US-09-759-5088-2	Sequence 57, Appl
876	22	57.9	872	10	US-09-815-242-12333	Sequence 12333, A	949	22	57.9	26926	9	US-09-866-565-57	Sequence 2, Appl
877	22	57.9	873	9	US-10-041-007-2	Sequence 2, Appl	950	22	56.6	1047	10	US-09-866-565-57	Sequence 2, Appl
878	22	57.9	873	9	US-10-041-007-33	Sequence 33, Appl	951	22	55.3	15	10	US-09-824-787B-89	Sequence 89, Appl
879	22	57.9	874	10	US-09-765-111A-6	Sequence 6, Appl	952	22	55.3	15	10	US-09-005-243-71	Sequence 71, Appl
880	22	57.9	877	10	US-09-746-491-47	Sequence 47, Appl	953	22	55.3	15	10	US-09-924-683-71	Sequence 71, Appl
881	22	57.9	890	10	US-09-815-242-10314	Sequence 10314, A	954	22	55.3	17	10	US-09-966-871-52	Sequence 52, Appl
882	22	57.9	892	10	US-09-815-242-13765	Sequence 13765, A	955	22	55.3	17	12	US-10-039-645-52	Sequence 52, Appl
883	22	57.9	892	10	US-09-815-242-11148	Sequence 11148, A	956	22	55.3	27	10	US-09-864-761-34137	Sequence 34137, A
884	22	57.9	901	10	US-09-815-242-13901	Sequence 13901, A	957	22	55.3	34	10	US-09-864-761-45367	Sequence 45367, A
885	22	57.9	909	12	US-10-078-929-168	Sequence 168, App	958	22	55.3	40	10	US-09-925-302-848	Sequence 848, App
886	22	57.9	917	9	US-09-801-574-90	Sequence 90, Appl	959	22	55.3	40	10	US-09-864-761-36532	Sequence 36532, A
887	22	57.9	935	9	US-10-029-217A-31	Sequence 31, Appl	960	22	55.3	42	10	US-09-864-761-45400	Sequence 45400, A
888	22	57.9	935	9	US-10-029-217A-32	Sequence 32, Appl	961	22	55.3	44	10	US-09-864-761-45400	Sequence 45400, A
889	22	57.9	937	10	US-09-949-192-19	Sequence 19, Appl	962	22	55.3	46	10	US-09-973-404-1	Sequence 1, Appl
890	22	57.9	938	9	US-10-028-217A-28	Sequence 28, Appl	963	22	55.3	48	10	US-09-864-761-35521	Sequence 35521, A
891	22	57.9	938	9	US-10-029-217A-29	Sequence 29, Appl	964	22	55.3	54	10	US-09-864-761-35507	Sequence 35507, A
892	22	57.9	946	10	US-09-741-669-470	Sequence 470, App	965	22	55.3	54	10	US-09-867-550-18	Sequence 18, Appl
893	22	57.9	946	10	US-09-912-020-326	Sequence 326, App	966	22	55.3	58	10	US-09-864-761-43518	Sequence 43518, A
894	22	57.9	957	10	US-09-801-574-80	Sequence 80, Appl	967	22	55.3	58	10	US-09-864-761-46000	Sequence 46000, A
895	22	57.9	969	8	US-08-945-749-3	Sequence 3, Appl	968	22	55.3	60	10	US-09-864-761-37408	Sequence 37408, A

```

969 21 55.3 61 10 US-09-864-761-44746 Sequence 44746, A
970 21 55.3 65 8 US-08-808-031A-16 Sequence 16, Appl
971 21 55.3 65 8 US-08-808-031A-16 Sequence 17, Appl
972 21 55.3 67 10 US-09-864-761-45319 Sequence 45319, A
973 21 55.3 68 10 US-09-867-550-378 Sequence 378, App
974 21 55.3 71 10 US-09-864-761-39582 Sequence 39582, A
975 21 55.3 71 10 US-09-864-761-40657 Sequence 40657, A
976 21 55.3 71 10 US-09-734-569-142 Sequence 142, App
977 21 55.3 73 9 US-09-861-876-139 Sequence 139, App
978 21 55.3 73 10 US-09-864-761-38358 Sequence 38358, A
979 21 55.3 76 10 US-09-925-299-1081 Sequence 1081, App
980 21 55.3 77 10 US-09-925-302-608 Sequence 608, App
981 21 55.3 79 10 US-09-952-768-65 Sequence 65, Appl
982 21 55.3 84 12 US-10-067-989-10 Sequence 10, Appl
983 21 55.3 86 10 US-09-864-761-34597 Sequence 34597, A
984 21 55.3 86 10 US-09-987-108-13 Sequence 12, Appl
985 21 55.3 86 10 US-09-987-108-13 Sequence 13, Appl
986 21 55.3 89 10 US-09-728-721-32 Sequence 32, Appl
987 21 55.3 89 10 US-09-881-752A-30 Sequence 10, Appl
988 21 55.3 89 10 US-09-931-071-10 Sequence 32, Appl
989 21 55.3 89 12 US-10-105-931-32 Sequence 7, Appl
990 21 55.3 90 10 US-09-841-739-7 Sequence 46509, A
991 21 55.3 90 10 US-09-864-761-46509 Sequence 6, Appl
992 21 55.3 90 10 US-09-767-215-6 Sequence 45, Appl
993 21 55.3 90 10 US-09-925-300-1777 Sequence 1777, App
994 21 55.3 91 10 US-09-864-761-33616 Sequence 33616, A
995 21 55.3 91 10 US-09-864-761-33616 Sequence 34143, A
996 21 55.3 92 12 US-10-014-269-29 Sequence 29, Appl
997 21 55.3 92 12 US-10-014-269-29 Sequence 98, App
998 21 55.3 93 10 US-09-764-869-988 Sequence 29, Appl
999 21 55.3 96 10 US-09-798-029-29 Sequence 86, Appl
1000 21 55.3 97 9 US-09-864-921-86

```

## ALIGNMENTS

```

RESULT 1
US-09-764-847-543
: Sequence 543, Application US/09764847
: Patent No. US20020132767A1
: GENERAL INFORMATION:
: APPLICANT: Rosen et al.
: TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
: FILE REFERENCE: PC009
: CURRENT APPLICATION NUMBER: US/09/764,847
: PRIOR FILING DATE: 2001-01-17
: NUMBER OF SEQ ID NOS: 2003
: SOFTWARE: Patentln Ver. 2.0
: SEQ ID NO 543
: LENGTH: 94
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: SITE
: LOCATION: (86)
: OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
: NAME/KEY: SITE
: LOCATION: (94)
: OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-847-543

```

```

Query Match          78.9%; Score 30; DB 10; Length 94;
Best Local Similarity 75.0%; Pred. No. 9.5;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

OY 2 LXDLEA 9
    |||||
Db 21 LSDLES 28

```

RESULT 2

```

US-10-014-717-3
: Sequence 3, Application US/10014717
: Publication No. US20020192778A1
: GENERAL INFORMATION:
: APPLICANT: Schnupp, Thomas
: APPLICANT: Ligon, James
: APPLICANT: Molnar, Istvan
: APPLICANT: Zirkle, Ross
: APPLICANT: Cyr, Devon
: APPLICANT: Goerlach, Joern
: TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
: FILE REFERENCE: 4-30582A
: CURRENT APPLICATION NUMBER: US/10/014,717
: PRIOR FILING DATE: 2001-11-13
: NUMBER OF SEQ ID NOS: 30
: SOFTWARE: Patentln Ver. 2.0
: SEQ ID NO 3
: LENGTH: 1410
: TYPE: PRT
: ORGANISM: Sorangium cellulosum
US-10-014-717-3

```

```

Query Match          78.9%; Score 30; DB 9; Length 1410;
Best Local Similarity 75.0%; Pred. No. 1.8e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

OY 2 LXDLEA 9
    |||||
Db 474 LDDMLEA 481

```

```

RESULT 3
US-09-989-920-200
: Sequence 200, Application US/09989920
: Patent No. US20020172957A1
: GENERAL INFORMATION:
: APPLICANT: Macina, Roberto
: APPLICANT: Recipon, Hervé
: APPLICANT: Chen, Sel-yu
: APPLICANT: Sun, Yonming
: APPLICANT: Liu, Chenghua
: TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and P
: FILE REFERENCE: DEX-0291
: CURRENT APPLICATION NUMBER: US/09/989,920
: PRIOR FILING DATE: 2001-11-21
: PRIOR APPLICATION NUMBER: 60/252,500
: NUMBER OF SEQ ID NOS: 284
: SOFTWARE: Patentln version 3.1
: SEQ ID NO 200
: LENGTH: 594
: TYPE: PRT
: ORGANISM: Homo sapien
US-09-989-920-200

```

```

Query Match          76.3%; Score 29; DB 9; Length 594;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 4 DDLLEA 9
    |||||
Db 216 DDLLEA 221

```

```

RESULT 4
US-09-989-920-249
: Sequence 249, Application US/09989920
: Patent No. US20020172957A1
: GENERAL INFORMATION:
: APPLICANT: Macina, Roberto
: APPLICANT: Recipon, Hervé

```

```

; APPLICANT: Chen, Sei-Yu
; APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chenghua
; TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and Prot
; FILE REFERENCE: DEX-0291
; CURRENT APPLICATION NUMBER: US/09/989,920
; CURRENT FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: 60/252,500
; PRIOR FILING DATE: 2000-11-22
; NUMBER OF SEQ ID NOS: 284
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 249
; LENGTH: 594
; TYPE: PRT
; ORGANISM: Homo sapien
; US-09-989-920-249

Query Match          76.3%; Score 29; DB 9; Length 594;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      4 DDLLEA 9
        |||||
Db      216 DDLLEA 221

RESULT 5
US-09-934-899-2
; Sequence 2, Application US/09934899
; Patent No. US20020102697A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Siqun
; APPLICANT: Wang, Tao
; APPLICANT: Koffas, Mattheos
; APPLICANT: Odom, J. Martin
; APPLICANT: Ye, Rick
; TITLE OF INVENTION: Genes encoding exopolysaccharide production
; FILE REFERENCE: CL1633 US NA
; CURRENT APPLICATION NUMBER: US/09/934,899
; CURRENT FILING DATE: 2001-08-22
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Microsoft Office 97
; SEQ ID NO: 2
; LENGTH: 293
; TYPE: PRT
; ORGANISM: Methylobionas 16a
; US-09-934-899-2

Query Match          73.7%; Score 28; DB 10; Length 293;
Best Local Similarity 62.5%; Pred. No. 83;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      1 VLXDDLLE 8
        :| |||:|
Db      131 ILPDDLLE 138

RESULT 6
US-09-934-868-22
; Sequence 22, Application US/09934868
; Patent No. US20020137190A1
; GENERAL INFORMATION:
; APPLICANT: Koffas, Mattheos
; APPLICANT: Odom, James M
; APPLICANT: Schenle, Andreas J
; TITLE OF INVENTION: DENITRIFYING METHANOTROPHIC BACTERIAL STRAIN
; FILE REFERENCE: CL1596 US NA
; CURRENT APPLICATION NUMBER: US/09/934,868
; CURRENT FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/229,858
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: Microsoft Office 97

```

```

; SEQ ID NO: 22
; LENGTH: 293
; TYPE: PRT
; ORGANISM: Methylobionas 16a
; US-09-934-868-22

Query Match          73.7%; Score 28; DB 10; Length 293;
Best Local Similarity 62.5%; Pred. No. 83;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      1 VLXDDLLE 8
        :| |||:|
Db      131 ILPDDLLE 138

RESULT 7
US-09-815-242-13353
; Sequence 13353, Application US/09815242
; Patent No. US2002061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlson, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 13353
; LENGTH: 296
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; US-09-815-242-13353

Query Match          73.7%; Score 28; DB 10; Length 296;
Best Local Similarity 66.7%; Pred. No. 84;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      1 VLXDDLLEA 9
        :| |||||
Db      266 VILDDLLEA 274

RESULT 8
US-09-815-242-13677
; Sequence 13677, Application US/09815242
; Patent No. US2002061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlson, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel

```

```
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: US/09/815,242
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 13677
LENGTH: 296
TYPE: PRT
ORGANISM: Streptococcus pneumoniae
US-09-815-242-13677
```

```
Query Match 73.7% Score 28; DB 10; Length 296;
Best Local Similarity 66.7%; Pred. No. 84;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 1 LKDDLEA 9
1: |||||
Db 266 VILQDLLEA 274
```

```
RESULT 9
US-09-864-761-41433
Sequence 41433, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecm1ca-X-1
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
```

```
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1
SEQ ID NO 41433
LENGTH: 63
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC013612.3
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.3
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.9
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.4
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.6
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.8
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.1
OTHER INFORMATION: EXPRESSED IN HEPA, SIGNAL = 3
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.1
OTHER INFORMATION: SWISSPROT HIT: Q15034, EVALUATE 7.00e-05
OTHER INFORMATION: EST_HUMAN HIT: AW963676.1, EVALUATE 1.00e-31
US-09-864-761-41433
```

```
Query Match 71.1% Score 27; DB 10; Length 63;
Best Local Similarity 75.0%; Pred. No. 25;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 2 LKDDLEA 9
1: |||||
Db 5 LEDDLRA 12
RESULT 10
US-10-052-586-380
Sequence 380, Application US/10052586
Patent No. US20020127584A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3450R1C1
CURRENT FILING DATE: 2002-01-15
PRIOR APPLICATION NUMBER: 60/059263
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/059266
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
```



```
; PRIOR APPLICATION NUMBER: 60/088811
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088824
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088825
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088826
; PRIOR FILING DATE: 1998-06-10
; PRIOR APPLICATION NUMBER: 60/088861
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088863
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/088876
; PRIOR FILING DATE: 1998-06-11
; PRIOR APPLICATION NUMBER: 60/089090
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089105
; PRIOR FILING DATE: 1998-06-12
; PRIOR APPLICATION NUMBER: 60/089512
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089514
; PRIOR FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089598
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089988
```

```
Query Match          71.1%; Score 27; DB 12; Length 140;
Best Local Similarity 62.5%; Pred. No. 60;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
```

```
OY      1 VLXDDLE 8
        |||||
Db      88 VLDDILQ 95
```

```
RESULT 11
US-09-815-242-10630
; Sequence 10630, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA 011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 10630
; LENGTH: 293
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-815-242-10630
```

```
Query Match          71.1%; Score 27; DB 10; Length 293;
Best Local Similarity 62.5%; Pred. No. 1,3e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
```

```
OY      2 LXDLDLEA 9
        :|||||
Db      215 IRDDILDA 222
```

```
RESULT 12
US-09-815-915-11
; Sequence 11, Application US/09815915
; Patent No. US20020025931A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Meyers, Rachel
; TITLE OF INVENTION: 3714, 16742, 23546, AND 13887 NOVEL
; FILE REFERENCE: 38155-20006.00
; CURRENT APPLICATION NUMBER: US/09/815,915
; CURRENT FILING DATE: 2001-03-23
; PRIOR APPLICATION NUMBER: US 60/191,846
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 419
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-815-915-11
```

```
Query Match          71.1%; Score 27; DB 10; Length 419;
Best Local Similarity 75.0%; Pred. No. 1,9e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
OY      1 VLXDDLE 8
        |||||
Db      328 VLDDDYLE 335
```

```
RESULT 13
US-09-952-013A-6
; Sequence 6, Application US/09952013A
; Patent No. US20020146801A1
; GENERAL INFORMATION:
; APPLICANT: GRUMET, Ingrid
; APPLICANT: VINGRON, Martin
; TITLE OF INVENTION: RNA POLYMERASE I TRANSCRIPTION FACTOR TIF-1A
; FILE REFERENCE: 38485-0007
; CURRENT APPLICATION NUMBER: US/09/952,013A
; CURRENT FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: PCT/DE00/00767
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: DE 199 11 992.9
; PRIOR FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 602
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-952-013A-6
```

```
Query Match          71.1%; Score 27; DB 10; Length 602;
Best Local Similarity 62.5%; Pred. No. 2,9e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
```

OY 2 LXDLEA 9  
: ||||:  
Db 484 IEDDLLES 491

RESULT 14  
US-09-735-927-2  
; Sequence 2, Application US/09735927  
; Patent No. US20020137128A1  
; GENERAL INFORMATION:  
; APPLICANT: WEI, Ming Hui et al.  
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,  
; TITLE OF INVENTION: AND USES THEREOF  
; FILE REFERENCE: CLO00834  
; CURRENT APPLICATION NUMBER: US/09/735,927  
; CURRENT FILING DATE: 2002-01-11  
; PRIOR APPLICATION NUMBER: 60/231,570  
; PRIOR FILING DATE: 2000-09-11  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 664  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-09-735-927-2

Query Match 71.1%; Score 27; DB 10; Length 664;  
Best Local Similarity 83.3%; Pred. No. 3.2e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 4 DDLEA 9  
: ||||:  
Db 554 DDLMEA 559

RESULT 15  
US-09-735-927-4  
; Sequence 4, Application US/09735927  
; Patent No. US20020137128A1  
; GENERAL INFORMATION:  
; APPLICANT: WEI, Ming Hui et al.  
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,  
; TITLE OF INVENTION: AND USES THEREOF  
; FILE REFERENCE: CLO00834  
; CURRENT APPLICATION NUMBER: US/09/735,927  
; CURRENT FILING DATE: 2002-01-11  
; PRIOR APPLICATION NUMBER: 60/231,570  
; PRIOR FILING DATE: 2000-09-11  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 664  
; TYPE: PRT  
; ORGANISM: Rabbit  
US-09-735-927-4

Query Match 71.1%; Score 27; DB 10; Length 664;  
Best Local Similarity 83.3%; Pred. No. 3.2e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 4 DDLEA 9  
: ||||:  
Db 554 DDLMEA 559

RESULT 16  
US-10-034-843-2  
; Sequence 2, Application US/10034843  
; Patent No. US2002011478A1  
; GENERAL INFORMATION:  
; APPLICANT: YU, Xuanchuan

APPLICANT: Miranda, Maricar  
; TITLE OF INVENTION: No. US2002011478A1 Human Ion Channel Protein and Polynucleo  
; FILE REFERENCE: LEX-0291-US  
; CURRENT APPLICATION NUMBER: US/10/034,843  
; CURRENT FILING DATE: 2001-12-27  
; PRIOR APPLICATION NUMBER: US 60/258,334  
; PRIOR FILING DATE: 2000-12-27  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 664  
; TYPE: PRT  
; ORGANISM: homo sapiens  
US-10-034-843-2

Query Match 71.1%; Score 27; DB 12; Length 664;  
Best Local Similarity 83.3%; Pred. No. 3.2e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 4 DDLEA 9  
: ||||:  
Db 554 DDLMEA 559

RESULT 17  
US-09-712-363-176  
; Sequence 176, Application US/09712363  
; Patent No. US20020164588A1  
; GENERAL INFORMATION:  
; APPLICANT: Eisenberg, David  
; APPLICANT: Rolsteijn, Sergio H.  
; APPLICANT: Marcotte, Edward M.  
; TITLE OF INVENTION: DETERMINING THE FUNCTIONS AND  
; TITLE OF INVENTION: INTERACTIONS OF PROTEINS BY COMPARATIVE ANALYSIS  
; FILE REFERENCE: 07419-032001  
; CURRENT APPLICATION NUMBER: US/09/712,363  
; CURRENT FILING DATE: 2000-11-13  
; PRIOR APPLICATION NUMBER: PCT/US00/02246  
; PRIOR FILING DATE: 2000-01-28  
; PRIOR APPLICATION NUMBER: 60/179,531  
; PRIOR FILING DATE: 2000-02-01  
; PRIOR APPLICATION NUMBER: 60/117,844  
; PRIOR FILING DATE: 1999-01-29  
; PRIOR APPLICATION NUMBER: 60/118,206,  
; PRIOR FILING DATE: 1999-02-01  
; PRIOR APPLICATION NUMBER: 60/126,593  
; PRIOR FILING DATE: 1999-03-26  
; PRIOR APPLICATION NUMBER: 60/134,093  
; PRIOR FILING DATE: 1999-05-14  
; PRIOR APPLICATION NUMBER: 60/134,092  
; PRIOR FILING DATE: 1999-05-14  
; PRIOR APPLICATION NUMBER: 60/165,124  
; PRIOR FILING DATE: 1999-11-12  
; PRIOR APPLICATION NUMBER: 60/165,086  
; PRIOR FILING DATE: 1999-11-12  
; NUMBER OF SEQ ID NOS: 292  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 176  
; LENGTH: 1172  
; TYPE: PRT  
; ORGANISM: Mycobacterium tuberculosis  
US-09-712-363-176

Query Match 71.1%; Score 27; DB 9; Length 1172;  
Best Local Similarity 75.0%; Pred. No. 5.9e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 LXDLEA 9  
: ||||:  
Db 953 LPDLLEA 960

RESULT 18

```

US-09-843-676-141
; Sequence 141, Application US/09843676
; Patent No. US20020164786A1
GENERAL INFORMATION:
APPLICANT: Cecch, Thomas R.
; Lingner, Joachim
; Nakamura, Toru
; Chapman, Karen B.
; Morin, Gregg B.
; Harley, Calvin
; Andrews, William H.
TITLE OF INVENTION: No. US20020164786A1 Telomerase
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESSES:
ADDRESS: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/843,676
FILING DATE: 26-Apr-2001
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/854,050
FILING DATE: 09-MAY-1997
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-0029300S
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 141:
SEQUENCE CHARACTERISTICS:
LENGTH: 32 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: Linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..32
OTHER INFORMATION: /note="motif 4(B') peptide from
Saccharomyces cerevisiae EST2p"
SEQUENCE DESCRIPTION: SEQ ID NO: 141:
US-09-843-676-141

Query Match 68.4%; Score 26; DB 9; Length 32;
Best Local Similarity 62.5%; Pred. No. 19;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 VLXDDLE 8
; : : : : :
Db 20 LVYDDLE 27

```

```

GENERAL INFORMATION:
APPLICANT: Cecch, Thomas R.
; Lingner, Joachim
; Nakamura, Toru
; Chapman, Karen B.
; Morin, Gregg B.
; Harley, Calvin
; Andrews, William H.
TITLE OF INVENTION: No. US20020187471A1 Telomerase
NUMBER OF SEQUENCES: 171
CORRESPONDENCE ADDRESSES:
ADDRESS: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/766,253
FILING DATE: 19-Jan-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/846,017
FILING DATE: 1997-04-25
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-0029200S
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 141:
SEQUENCE CHARACTERISTICS:
LENGTH: 32 amino acids
TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: Linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..32
OTHER INFORMATION: /note="motif 4(B') peptide from
Saccharomyces cerevisiae EST2p"
SEQUENCE DESCRIPTION: SEQ ID NO: 141:
US-09-766-253-141

Query Match 68.4%; Score 26; DB 9; Length 32;
Best Local Similarity 62.5%; Pred. No. 19;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 VLXDDLE 8
; : : : : :
Db 20 LVYDDLE 27

RESULT 20
US-09-823-266-19
; Sequence 19, Application US/09823266
; Patent No. US20020127688A1
GENERAL INFORMATION:
APPLICANT: Burgess, Richard
; Applicant: Arthur, Terrance
; Applicant: Anthony, Larry
; Applicant: Bergendahl, Vlet
; Applicant: Pietz, Bradley
TITLE OF INVENTION: Sigma binding region of RNA polymerase and uses thereof

```



```
FILE REFERENCE: 800.025051
CURRENT APPLICATION NUMBER: US/09/823.266
CURRENT FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: US 60/193.116
NUMBER OF SEQ ID NOS: 26
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 19
LENGTH: 37
TYPE: PRT
ORGANISM: Escherichia coli
US-09-823-266-19

Query Match
Best Local Similarity 68.4%; Score 26; DB 10; Length 37;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 4 DDLLEA 9
Db 23 DDLLEA 28

RESULT 21
US-09-864-761-36279
Sequence 36279, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecmics-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180.312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
```

```
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 36279
LENGTH: 45
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC007114.7
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1
OTHER INFORMATION: EXPRESSED IN HEL100, SIGNAL = 1
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.94
OTHER INFORMATION: SWISSPROT HIT: P21529, EVALU 4.70e-02
OTHER INFORMATION: EST_HUMAN HIT: AV705451.1, EVALU 3.00e-18
US-09-864-761-36279

Query Match
Best Local Similarity 62.5%; Score 26; DB 10; Length 45;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Oy 2 LXDLEA 9
Db 20 LXDLEA 27

RESULT 22
US-09-843-676-124
Sequence 124, Application US/09843676
Patent No. US20020164786A1
GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
APPLICANT: Lingner, Joachim
APPLICANT: Nakamura, Toru
APPLICANT: Chapman, Karen B.
APPLICANT: Morin, Gregg B.
APPLICANT: Harley, Calvin
APPLICANT: Andrews, William H.
TITLE OF INVENTION: No. US20020164786A1 Telomerase
NUMBER OF SEQUENCES: 225
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/843,676
FILING DATE: 26-Apr-2001
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/854,050
FILING DATE: 09-MAY-1997
APPLICATION NUMBER: US 08/846,017
FILING DATE: 25-APR-1997
APPLICATION NUMBER: US 08/844,419
FILING DATE: 18-APR-1997
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-00293005
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 124:
SEQUENCE CHARACTERISTICS:
LENGTH: 49 amino acids
```

TYPE: amino acid  
STRANDEDNESS: <Unknown>  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 124:  
US-09-843-676-124

Query Match 68.4%; Score 26; DB 9; Length 49;  
Best Local Similarity 62.5%; Pred. No. 31;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLXDDLE 8  
: : |||||  
Db 16 LVYDDLE 23

RESULT 23  
US-09-766-253-124  
Sequence 124, Application US/09766253  
Publication No. US20020187471A1

GENERAL INFORMATION:

APPLICANT: Cecch, Thomas R.

Lingner, Joachim

Nakamura, Toru

Chapman, Karen B.

Morin, Gregg B.

Harley, Calvin

Andrews, William H.

TITLE OF INVENTION: No. US20020187471a1e1 Telomerase

NUMBER OF SEQUENCES: 171

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP

STREET: Two Embarcadero Center, 8th Floor

CITY: San Francisco

STATE: California

COUNTRY: United States of America

ZIP: 94111

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/766,253

FILING DATE: 19-Jan-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/846,017

FILING DATE: 1997-04-25

APPLICATION NUMBER: US 08/724,643

FILING DATE: 01-OCT-1996

ATTORNEY/AGENT INFORMATION:

NAME: Apple, Randolph T.

REGISTRATION NUMBER: 36,429

REFERENCE/DOCKET NUMBER: 015389-00292005

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 576-0200

TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 124:

SEQUENCE CHARACTERISTICS:

LENGTH: 49 amino acids

TYPE: amino acid

STRANDEDNESS: <Unknown>

TOPOLOGY: linear

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 124:  
US-09-766-253-124

Query Match 68.4%; Score 26; DB 9; Length 49;  
Best Local Similarity 62.5%; Pred. No. 31;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLXDDLE 8

Db 16 LVYDDLE 23  
: : |||||

RESULT 24

US-09-843-676-17

Sequence 17, Application US/09843676

Patent No. US20020164786A1

GENERAL INFORMATION:

APPLICANT: Cecch, Thomas R.

Lingner, Joachim

Nakamura, Toru

Chapman, Karen B.

Morin, Gregg B.

Harley, Calvin

Andrews, William H.

TITLE OF INVENTION: No. US20020164786A1e1 Telomerase

NUMBER OF SEQUENCES: 225

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP

STREET: Two Embarcadero Center, 8th Floor

CITY: San Francisco

STATE: California

COUNTRY: United States of America

ZIP: 94111

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/843,676

FILING DATE: 26-Apr-2001

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/854,050

FILING DATE: 09-MAY-1997

APPLICATION NUMBER: US 08/846,017

FILING DATE: 25-APR-1997

APPLICATION NUMBER: US 08/844,419

FILING DATE: 18-APR-1997

APPLICATION NUMBER: US 08/724,643

FILING DATE: 01-OCT-1996

ATTORNEY/AGENT INFORMATION:

NAME: Apple, Randolph T.

REGISTRATION NUMBER: 36,429

REFERENCE/DOCKET NUMBER: 015389-00293005

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 576-0200

TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 17:

SEQUENCE CHARACTERISTICS:

LENGTH: 69 amino acids

TYPE: amino acid

STRANDEDNESS: No. US20020164786A1 Relevant

TOPOLOGY: No. US20020164786A1 Relevant

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 17:  
US-09-843-676-17

Query Match 68.4%; Score 26; DB 9; Length 69;  
Best Local Similarity 62.5%; Pred. No. 44;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLXDDLE 8  
: : |||||

Db 56 LVYDDLE 63

RESULT 25

US-09-766-253-17

Sequence 17, Application US/09766253

Publication No. US20020187471A1

```

GENERAL INFORMATION:
APPLICANT: Cech, Thomas R.
            Lingner, Joachim
            Nakamura, Toru
            Chapman, Karen B.
            Morin, Gregg B.
            Harley, Calvin
            Andrews, William H.
TITLE OF INVENTION: No. US20020187471A1el Telomerase
NUMBER OF SEQUENCES: 171
CORRESPONDENCE ADDRESS:
ADDRESSER: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: United States of America
ZIP: 94111

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/766,253
FILING DATE: 19-Jan-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/846,017
FILING DATE: 1997-04-25
APPLICATION NUMBER: US 08/724,643
FILING DATE: 01-OCT-1996
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 015389-002920US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 69 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 17:
US-09-766-253-17

Query Match          68.4%: Score 26; DB 9; Length 69;
Best Local Similarity 62.5%: Pred. No. 44;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLXDDL 8
   :: |||||
Db 56 LVYDDL 63

RESULT 26
US-09-867-550-10
; Sequence 10, Application US/09867550
; Patent No. US20020082206A1
GENERAL INFORMATION:
APPLICANT: Leach, Martin D.
APPLICANT: Mehrtan, Fuad.
APPLICANT: Conley, Pamela
APPLICANT: Law, Debbie
APPLICANT: Topper, James
TITLE OF INVENTION: No. US20020082206A1el Polynucleotides from Atherogenic Cells and
TITLE OF INVENTION: Thereby
FILE REFERENCE: 21402-013 (Cura-313)
CURRENT APPLICATION NUMBER: US/09/867,550
CURRENT FILING DATE: 2001-09-20
PRIOR APPLICATION NUMBER: USSN 60/208,427

```

```

PRIOR FILING DATE: 2000-05-30
NUMBER OF SEQ ID NOS: 2125
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10
LENGTH: 148
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: VARIANT
LOCATION: (1)
OTHER INFORMATION: wherein Xaa may be any one of Ala or Arg or Glu or Gln or Gly
US-09-867-550-10

Query Match          68.4%: Score 26; DB 10; Length 148;
Best Local Similarity 71.4%: Pred. No. 1e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VLXDDL 7
   :: |||||
Db 46 ILGDDL 52

RESULT 27
US-09-925-302-599
; Sequence 599, Application US/09925302
; Patent No. US20020044941A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA104
CURRENT APPLICATION NUMBER: US/09/925,302
PRIOR FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05918
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 896
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 599
LENGTH: 151
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (56)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-302-599

Query Match          68.4%: Score 26; DB 10; Length 151;
Best Local Similarity 83.3%: Pred. No. 1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 DDLEA 9
   |||||
Db 87 DDLES 92

RESULT 28
US-09-925-297-793
; Sequence 793, Application US/09925297
; Patent No. US20020081659A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PA105
CURRENT APPLICATION NUMBER: US/09/925,297
CURRENT FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05989
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 928

```

```
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 793
; LENGTH: 190
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (3)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-297-793

Query Match
Best Local Similarity 68.4%; Score 26; DB 10; Length 190;
Pred. No. 1.3e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 2 LXDLLA 9
; :|:|:|:|
Db 90 LVDELLA 97

RESULT 29
US-09-917-974-2
; Sequence 2, Application US/09917974
; Patent No. US20020042389A1
; GENERAL INFORMATION:
; APPLICANT: Ligensa, Tanja
; APPLICANT: Schmacher, Ralf
; APPLICANT: Weidner, Michael
; TITLE OF INVENTION: IGF-1 Receptor Interacting Proteins
; FILE REFERENCE: 09/453,195
; CURRENT APPLICATION NUMBER: US/09/917,974
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 09/453,195
; PRIOR FILING DATE: 1999-12-02
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 333
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Xaa at position 42, 47, and 48 is any one of the twenty naturally
; OTHER INFORMATION: occurring amino acids.
US-09-917-974-2

Query Match
Best Local Similarity 68.4%; Score 26; DB 10; Length 333;
Pred. No. 2.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 4 DDLLEA 9
; :|:|:|:|
Db 269 DDLLES 274

RESULT 30
US-10-013-056-2
; Sequence 2, Application US/10013056
; Patent No. US20020146728A1
; GENERAL INFORMATION:
; APPLICANT: Ligensa, Tanja
; APPLICANT: Schmacher, Ralf
; APPLICANT: Weidner, Michael
; TITLE OF INVENTION: IGF-1 Receptor Interacting Proteins
; FILE REFERENCE: 09/453,195
; CURRENT APPLICATION NUMBER: US/10/013,056
; PRIOR FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: EPO 98122992.5
; PRIOR FILING DATE: 1998-12-03
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 333
; TYPE: PRT
```

```
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Xaa at position 42, 47, and 48 is any one of the twenty natura
; OTHER INFORMATION: occurring amino
; OTHER INFORMATION: acids.
US-10-013-056-2

Query Match
Best Local Similarity 68.4%; Score 26; DB 12; Length 333;
Pred. No. 2.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 4 DDLLEA 9
; :|:|:|:|
Db 269 DDLLES 274

RESULT 31
US-09-754-446-4
; Sequence 4, Application US/09754446
; Patent No. US20010016334A1
; GENERAL INFORMATION:
; APPLICANT: Wallis, Nicola G.
; TITLE OF INVENTION: MuRF
; FILE REFERENCE: GMI0097
; CURRENT APPLICATION NUMBER: US/09/754,446
; CURRENT FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 09/143,954
; PRIOR FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 448
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-754-446-4

Query Match
Best Local Similarity 68.4%; Score 26; DB 10; Length 448;
Pred. No. 3.3e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 VLXDLLA 9
; :|:|:|:|
Db 69 ILVDVILTA 77

RESULT 32
US-09-905-291A-255
; Sequence 255, Application US/09905291A
; Patent No. US20020160374A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Geo, Wei-Oiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gertlisen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gunney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kijavlin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
```

```
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William, I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: 10466-14
CURRENT FILING DATE: 2001-07-12
PRIOR APPLICATION NUMBER: PCT/US00/04414
PRIOR FILING DATE: 2000-02-22
PRIOR APPLICATION NUMBER: US 60/143,048
PRIOR FILING DATE: 1999-07-07
PRIOR APPLICATION NUMBER: US 60/145,698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: US 60/146,222
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: PCT/US99/20534
PRIOR FILING DATE: 1999-09-08
PRIOR APPLICATION NUMBER: PCT/US99/20944
PRIOR FILING DATE: 1999-09-13
PRIOR APPLICATION NUMBER: PCT/US99/21090
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/21547
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/23089
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: PCT/US99/28214
PRIOR FILING DATE: 1999-11-29
PRIOR APPLICATION NUMBER: PCT/US99/28313
PRIOR FILING DATE: 1999-11-30
PRIOR APPLICATION NUMBER: PCT/US99/28564
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/28565
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/30095
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: PCT/US99/30911
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US99/30999
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US00/00219
NUMBER OF SEQ ID NOS: 423
SEQ ID NO 255
LENGTH: 452
TYPE: PRT
ORGANISM: Homo Sapien
US-09-905-291A-255

Query Match      68.4%; Score 26; DB 9; Length 452;
Best Local Similarity 62.5%; Pred. No. 3.4e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY      2 LXDILLEA 9
Db      352 IVDLELEA 359
```

```
RESULT 33
US-09-902-853-255
; Sequence 255, Application US/09902853
; Publication No. US20020192659A1
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerltsen, Mary E.
```

```
APPLICANT: Goddard, A.
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth, J.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Mather, Jennie P.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William, I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: 10466-14
CURRENT FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: US/09/902,853
PRIOR FILING DATE: 2000-09-18
PRIOR APPLICATION NUMBER: US 60/143,048
PRIOR FILING DATE: 1999-07-07
PRIOR APPLICATION NUMBER: US 60/145,698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: PCT/US99/20594
PRIOR FILING DATE: 1999-09-08
PRIOR APPLICATION NUMBER: PCT/US99/20944
PRIOR FILING DATE: 1999-09-13
PRIOR APPLICATION NUMBER: PCT/US99/21090
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/21547
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/23089
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: PCT/US99/28214
PRIOR FILING DATE: 1999-11-29
PRIOR APPLICATION NUMBER: PCT/US99/28313
PRIOR FILING DATE: 1999-11-30
PRIOR APPLICATION NUMBER: PCT/US99/28564
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/28565
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/30095
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: PCT/US99/30911
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US99/30999
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US00/00219
NUMBER OF SEQ ID NOS: 423
SEQ ID NO 255
LENGTH: 452
TYPE: PRT
ORGANISM: Homo Sapien
US-09-902-853-255

Query Match      68.4%; Score 26; DB 9; Length 452;
Best Local Similarity 62.5%; Pred. No. 3.4e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY      2 LXDILLEA 9
Db      352 IVDLELEA 359
```

```
RESULT 34
US-09-909-320-255
; Sequence 255, Application US/09909320
; Patent No. US20020132240A1
```

GENERAL INFORMATION:  
APPLICANT: Genentech, Inc.  
APPLICANT: Ashkenazi, Avi  
APPLICANT: Botstein, David  
APPLICANT: Desnoyers, Luc  
APPLICANT: Eaton, Dan L.  
APPLICANT: Ferrara, Napoleone  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Fong, Sherman  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerber, Hanspeter  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, A.  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Hillan, Kenneth, J.  
APPLICANT: Kijavlin, Ivar J.  
APPLICANT: Mather, Jennie P.  
APPLICANT: Pan, James  
APPLICANT: Paoni, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William, I.  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE OF INVENTION: Acids Encoding the Same  
FILE REFERENCE: 10466-14  
CURRENT APPLICATION NUMBER: US/09/909,320  
CURRENT FILING DATE: 2002-01-04  
PRIOR APPLICATION NUMBER: PCT/US00/04414  
PRIOR FILING DATE: 2000-02-22  
PRIOR APPLICATION NUMBER: US 60/143,048  
PRIOR FILING DATE: 1999-07-07  
PRIOR APPLICATION NUMBER: US 60/145,698  
PRIOR FILING DATE: 1999-07-26  
PRIOR APPLICATION NUMBER: US 60/146,222  
PRIOR FILING DATE: 1999-07-28  
PRIOR APPLICATION NUMBER: PCT/US99/20594  
PRIOR FILING DATE: 1999-09-08  
PRIOR APPLICATION NUMBER: PCT/US99/20944  
PRIOR FILING DATE: 1999-09-13  
PRIOR APPLICATION NUMBER: PCT/US99/21090  
PRIOR FILING DATE: 1999-09-15  
PRIOR APPLICATION NUMBER: PCT/US99/21547  
PRIOR FILING DATE: 1999-09-15  
PRIOR APPLICATION NUMBER: PCT/US99/23089  
PRIOR FILING DATE: 1999-10-05  
PRIOR APPLICATION NUMBER: PCT/US99/28214  
PRIOR FILING DATE: 1999-11-29  
PRIOR APPLICATION NUMBER: PCT/US99/28313  
PRIOR FILING DATE: 1999-11-30  
PRIOR APPLICATION NUMBER: PCT/US99/28564  
PRIOR FILING DATE: 1999-12-02  
PRIOR APPLICATION NUMBER: PCT/US99/28565  
PRIOR FILING DATE: 1999-12-02  
PRIOR APPLICATION NUMBER: PCT/US99/30095  
PRIOR FILING DATE: 1999-12-16  
PRIOR APPLICATION NUMBER: PCT/US99/30911  
PRIOR FILING DATE: 1999-12-20  
PRIOR APPLICATION NUMBER: PCT/US99/30999  
PRIOR FILING DATE: 1999-12-20  
PRIOR APPLICATION NUMBER: PCT/US00/00219  
PRIOR FILING DATE: 2000-01-05  
NUMBER OF SEQ ID NOS: 423  
SEQ ID NO 255  
LENGTH: 452  
TYPE: PRT  
ORGANISM: Homo Sapien  
US-09-909-320-255

Query Match 68.4%; Score 26; DB 10; Length 452;

Best Local Similarity 62.5%; Pred. No. 3.4e+02;  
Matches 5: Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
QY 2 LXDLLLA 9  
: | : | | | |  
Db 352 IVDLLLEA 359  
RESULT 35  
US-09-909-088B-255  
Sequence 255, Application US/09909088B  
Patent No. US20020146709A1  
GENERAL INFORMATION:  
APPLICANT: Genentech, Inc.  
APPLICANT: Ashkenazi, Avi  
APPLICANT: Botstein, David  
APPLICANT: Desnoyers, Luc  
APPLICANT: Eaton, Dan L.  
APPLICANT: Ferrara, Napoleone  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Fong, Sherman  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerber, Hanspeter  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, A.  
APPLICANT: Godowski, Paul J.  
APPLICANT: Grimaldi, Christopher J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Hillan, Kenneth, J.  
APPLICANT: Kijavlin, Ivar J.  
APPLICANT: Mather, Jennie P.  
APPLICANT: Pan, James  
APPLICANT: Paoni, Nicholas F.  
APPLICANT: Roy, Margaret Ann  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Williams, P. Mickey  
APPLICANT: Wood, William, I.  
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
FILE OF INVENTION: Acids Encoding the Same  
FILE REFERENCE: 10466-14  
CURRENT APPLICATION NUMBER: US/09/909,088B  
CURRENT FILING DATE: 2001-07-18  
PRIOR APPLICATION NUMBER: PCT/US00/04414  
PRIOR FILING DATE: 2000-02-22  
PRIOR APPLICATION NUMBER: US 60/143,048  
PRIOR FILING DATE: 1999-07-07  
PRIOR APPLICATION NUMBER: US 60/145,698  
PRIOR FILING DATE: 1999-07-26  
PRIOR APPLICATION NUMBER: US 60/146,222  
PRIOR FILING DATE: 1999-07-28  
PRIOR APPLICATION NUMBER: PCT/US99/20594  
PRIOR FILING DATE: 1999-09-08  
PRIOR APPLICATION NUMBER: PCT/US99/20944  
PRIOR FILING DATE: 1999-09-13  
PRIOR APPLICATION NUMBER: PCT/US99/21090  
PRIOR FILING DATE: 1999-09-15  
PRIOR APPLICATION NUMBER: PCT/US99/21547  
PRIOR FILING DATE: 1999-09-15  
PRIOR APPLICATION NUMBER: PCT/US99/23089  
PRIOR FILING DATE: 1999-10-05  
PRIOR APPLICATION NUMBER: PCT/US99/28214  
PRIOR FILING DATE: 1999-11-29  
PRIOR APPLICATION NUMBER: PCT/US99/28313  
PRIOR FILING DATE: 1999-11-30  
PRIOR APPLICATION NUMBER: PCT/US99/28564  
PRIOR FILING DATE: 1999-12-02  
PRIOR APPLICATION NUMBER: PCT/US99/28565  
PRIOR FILING DATE: 1999-12-02  
PRIOR APPLICATION NUMBER: PCT/US99/30095  
PRIOR FILING DATE: 1999-12-16  
PRIOR APPLICATION NUMBER: PCT/US99/30911  
PRIOR FILING DATE: 1999-12-20

```

; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 255
; LENGTH: 452
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-909-088B-255

Query Match
Best Local Similarity 68.4%; Score 26; DB 10; Length 452;
Pred. No. 3.4e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 LXXDLEA 9
: |||||
Db 352 IVDLEA 359

RESULT 36
US-09-754-446-2
; Sequence 2; Application US/0975446
; Patent No. US20010016334A1
; GENERAL INFORMATION:
; APPLICANT: Wallis, Nicola G.
; TITLE OF INVENTION: MuF
; FILE REFERENCE: GM10097
; CURRENT APPLICATION NUMBER: US/09/754,446
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 09/143,954
; PRIOR FILING DATE: 1998-08-31
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 457
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-754-446-2

Query Match
Best Local Similarity 68.4%; Score 26; DB 10; Length 457;
Pred. No. 3.4e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 VLXDLEA 9
: |||||
Db 78 ILVDVLEA 86

RESULT 37
US-09-815-242-13478
; Sequence 13478; Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
```

```

; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13478
; LENGTH: 457
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-815-242-13478

Query Match
Best Local Similarity 68.4%; Score 26; DB 10; Length 457;
Pred. No. 3.4e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 VLXDLEA 9
: |||||
Db 78 ILVDVLEA 86

RESULT 38
US-09-815-242-5140
; Sequence 5140; Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5140
; LENGTH: 568
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-815-242-5140

Query Match
Best Local Similarity 68.4%; Score 26; DB 10; Length 568;
Pred. No. 4.3e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 VLXDLEA 9
: |||||
Db 105 VLDDIGE 113
```

```

RESULT 39
US-10-071-751-30
; Sequence 30, Application US/10071751
; Patent No. US20020142352A1
GENERAL INFORMATION:
APPLICANT: Hunter, Shirley Wu
Sim, Gek-Ke
Webster, Eric R.
TITLE OF INVENTION: NOVEL ECTOPARASITE SALIVA PROTEINS AND
APPARATUS TO COLLECT SUCH PROTEINS
NUMBER OF SEQUENCES: 88
CORRESPONDENCE ADDRESS:
ADDRESSER: SHERIDAN ROSS P.C.
STREET: 1560 BROADWAY, SUITE 1200
CITY: DENVER
STATE: CO
COUNTRY: U.S.A.
ZIP: 80202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/071,751
FILING DATE: 07-Feb-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/171,156
FILING DATE: 1998-10-09
ATTORNEY/AGENT INFORMATION:
NAME: Connell, Gary J
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2618-17-C4-PUS
TELECOMMUNICATION INFORMATION:
TELEPHONE: 303/863-9700
TELEFAX: 303/863-0223
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 586 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Xaa = any amino acid
LOCATION: 379
SEQUENCE DESCRIPTION: SEQ ID NO: 30:
US-10-071-751-30
Query Match 68.4%; Score 26; DB 12; Length 586;
Best Local Similarity 44.4%; Pred. No. 4.5e+02;
Matches 4; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
OY 1 VLXDDLEA 9
: | | | | |
Db 429 ILIDELMDA 437

```

```

STATE: MA
COUNTRY: US
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: Windows95
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/115,178
FILING DATE: 02-Apr-2002
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/819,177
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Fasse, J. Peter
REGISTRATION NUMBER: 32,983
REFERENCE/DOCKET NUMBER: 07917/037001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-542-5070
TELEFAX: 617-542-8906
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 660 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
FEATURE:
OTHER INFORMATION: J1P-1 protein
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-115-178-1
Query Match 68.4%; Score 26; DB 12; Length 660;
Best Local Similarity 66.7%; Pred. No. 5.1e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
OY 1 VLXDDLEA 9
: | | | | |
Db 571 VRADDALEA 579

```

```

RESULT 41
US-09-767-870-8
; Sequence 8, Application US/09767870
; Patent No. US20020037549A1
GENERAL INFORMATION:
APPLICANT: Ruben et al.
TITLE OF INVENTION: ABC Transport Polynucleotides, Polypeptides, and Antibodies
FILE REFERENCE: P1010P1
CURRENT APPLICATION NUMBER: US/09/767,870
CURRENT FILING DATE: 2001-01-24
PRIOR APPLICATION NUMBER: PCT/US00/19736
PRIOR FILING DATE: 2000-07-20
PRIOR APPLICATION NUMBER: 60/145,215
PRIOR FILING DATE: 1999-07-23
PRIOR APPLICATION NUMBER: 60/149,445
PRIOR FILING DATE: 1999-08-18
PRIOR APPLICATION NUMBER: 60/164,730
PRIOR FILING DATE: 1999-11-12
NUMBER OF SEQ ID NOS: 21
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 8
: | | | | |
LENGTH: 709
TYPE: PRT
ORGANISM: Homo sapiens
US-09-767-870-8
Query Match 68.4%; Score 26; DB 10; Length 709;
Best Local Similarity 83.3%; Pred. No. 5.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 4 DDLEA 9

```



```
Db      38  DDLVEA 43
||:|:|
RESULT 42
US-10-060-230-23
; Sequence 23, Application US/10060230
; Patent No. US20020173014A1
; GENERAL INFORMATION:
; APPLICANT: HILTONEN, Kaleervo
; TITLE OF INVENTION: A method to control cellular (3R)-hydroxyacyl-CoA
; TITLE OF INVENTION: esters, precursor molecules for polyhydroxyalkanoate
; FILE REFERENCE: 0365-0528P
; CURRENT APPLICATION NUMBER: US/10/060,230
; CURRENT FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: 19991667
; PRIOR FILING DATE: 1999-08-03
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 23
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-060-230-23

Query Match
Best Local Similarity 68.4%; Score 26; DB 9; Length 736;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY      1  VLXDDLEA 9
|:|:|:|
Db      204  VMPEDLVEA 212

RESULT 43
US-10-060-230-24
; Sequence 24, Application US/10060230
; Patent No. US20020173014A1
; GENERAL INFORMATION:
; APPLICANT: HILTONEN, Kaleervo
; TITLE OF INVENTION: A method to control cellular (3R)-hydroxyacyl-CoA
; TITLE OF INVENTION: esters, precursor molecules for polyhydroxyalkanoate
; FILE REFERENCE: 0365-0528P
; CURRENT APPLICATION NUMBER: US/10/060,230
; CURRENT FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: 19991667
; PRIOR FILING DATE: 1999-08-03
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 24
; LENGTH: 736
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-060-230-24

Query Match
Best Local Similarity 68.4%; Score 26; DB 9; Length 736;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY      1  VLXDDLEA 9
|:|:|:|
Db      204  VMPEDLVEA 212

RESULT 44
US-09-770-689A-5
; Sequence 5, Application US/09770689A
; Patent No. US20020115171A1
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE
; FILE REFERENCE: CL001079
; CURRENT APPLICATION NUMBER: US/09/770,689A
; CURRENT FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 780
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
US-09-770-689A-5

Query Match
Best Local Similarity 68.4%; Score 26; DB 10; Length 780;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      1  VLXDDLE 8
|:|:|:|
Db      24  VLSEDLQ 31

RESULT 45
US-09-770-689A-2
; Sequence 2, Application US/09770689A
; Patent No. US20020115171A1
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN RAS-LIKE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING THESE HUMAN RAS-LIKE
; FILE REFERENCE: CL001079
; CURRENT APPLICATION NUMBER: US/09/770,689A
; CURRENT FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 803
; TYPE: PRT
; ORGANISM: HUMAN
US-09-770-689A-2

Query Match
Best Local Similarity 68.4%; Score 26; DB 10; Length 803;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      1  VLXDDLE 8
|:|:|:|
Db      24  VLSEDLQ 31

RESULT 46
US-09-815-242-11216
; Sequence 11216, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlssen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
```

PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
NUMBER OF SEQ ID NOS: 14110  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 11216  
LENGTH: 829  
TYPE: PRT  
ORGANISM: Haemophilus influenzae  
US-09-815-242-11216

Query Match 68.4%; Score 26; DB 10; Length 829;  
Best Local Similarity 83.3%; Pred. No. 6.5e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 4 DDLLEA 9  
| | | | |  
Db 482 DDLLEA 487

RESULT 47  
US-10-041-007-4  
Sequence 4, Application US/10041007  
Patent No. US20020164736A1  
GENERAL INFORMATION:  
APPLICANT: Matsuda, Seichi P.T.  
TITLE OF INVENTION: Ginkgo Biloba Levopimaradiene Synthase  
FILE REFERENCE: P02081051  
CURRENT APPLICATION NUMBER: US/10/041,007  
CURRENT FILING DATE: 2002-01-07  
PRIOR APPLICATION NUMBER: US 60/259,881  
PRIOR FILING DATE: 2001-01-05  
NUMBER OF SEQ ID NOS: 41  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 4  
LENGTH: 868  
TYPE: PRT  
ORGANISM: Abies grandis  
US-10-041-007-4

Query Match 68.4%; Score 26; DB 9; Length 868;  
Best Local Similarity 55.6%; Pred. No. 6.8e+02;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 VLXDDLEA 9  
| | | | |  
Db 618 VLXDDLEA 626

RESULT 48  
US-09-887-586A-56  
Sequence 56, Application US/09887586A  
Patent No. US20020094556A1  
GENERAL INFORMATION:  
APPLICANT: Chappell, Joseph  
APPLICANT: No. US20020094556A1, Joseph P.  
APPLICANT: Starks, Courtney M.  
APPLICANT: Manna, Kathleen R.  
TITLE OF INVENTION: SYNTHASES  
FILE REFERENCE: 07678-025001  
CURRENT APPLICATION NUMBER: US/09/887,586A  
CURRENT FILING DATE: 2001-06-22  
PRIOR APPLICATION NUMBER: 09/398,395  
PRIOR FILING DATE: 1999-09-17  
PRIOR APPLICATION NUMBER: 60/130,628

PRIOR FILING DATE: 1999-04-22  
PRIOR APPLICATION NUMBER: 60/150,262  
PRIOR FILING DATE: 1999-08-23  
NUMBER OF SEQ ID NOS: 58  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 56  
LENGTH: 868  
TYPE: PRT  
ORGANISM: Abies grandis  
US-09-887-586A-56

Query Match 68.4%; Score 26; DB 10; Length 868;  
Best Local Similarity 55.6%; Pred. No. 6.8e+02;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 VLXDDLEA 9  
| | | | |  
Db 618 VLXDDLEA 626

RESULT 49  
US-09-903-012-56  
Sequence 56, Application US/09903012  
Patent No. US20020094557A1  
GENERAL INFORMATION:  
APPLICANT: Chappell, Joseph  
APPLICANT: No. US20020094557A1, Joseph P.  
APPLICANT: Starks, Courtney M.  
APPLICANT: Manna, Kathleen R.  
TITLE OF INVENTION: SYNTHASES  
FILE REFERENCE: 07678-025001  
CURRENT APPLICATION NUMBER: US/09/903,012  
CURRENT FILING DATE: 2001-07-11  
PRIOR APPLICATION NUMBER: 09/398,395  
PRIOR FILING DATE: 1999-09-17  
PRIOR APPLICATION NUMBER: 60/100,993  
PRIOR FILING DATE: 1998-09-18  
PRIOR APPLICATION NUMBER: 60/130,628  
PRIOR FILING DATE: 1999-04-22  
PRIOR APPLICATION NUMBER: 60/150,262  
PRIOR FILING DATE: 1999-08-23  
NUMBER OF SEQ ID NOS: 58  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 56  
LENGTH: 868  
TYPE: PRT  
ORGANISM: Abies grandis  
US-09-903-012-56

Query Match 68.4%; Score 26; DB 10; Length 868;  
Best Local Similarity 55.6%; Pred. No. 6.8e+02;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 VLXDDLEA 9  
| | | | |  
Db 618 VLXDDLEA 626

RESULT 50  
US-09-816-860A-2  
Sequence 2, Application US/09816860A  
Patent No. US20020081651A1  
GENERAL INFORMATION:  
APPLICANT: Meyers, Rachel  
APPLICANT: No. US20020081651A1 Human GTPase Activating Molecul  
TITLE OF INVENTION: Uses Therefor  
FILE REFERENCE: MNT-133  
CURRENT APPLICATION NUMBER: US/09/816,860A  
CURRENT FILING DATE: 2001-03-23  
PRIOR APPLICATION NUMBER: 60/191,859  
PRIOR FILING DATE: 2000-03-24  
NUMBER OF SEQ ID NOS: 3  
SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 2  
 LENGTH: 881  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 US-09-816-860A-2

Query Match 68.48; Score 26; DB 10; Length 881;  
 Best Local Similarity 62.58; Pred. No. 6.9e+02;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 VLXDDLLE 8  
 Db 24 VLSDDLQ 31

Search completed: December 27, 2002, 14:46:02  
 Job time : 12 secs

